Global Assimilation of Multi-Angle and Multi-Polarization SMOS Brightness Temperature Observations into the GEOS-5 Catchment Land Surface Model for Soil Moisture Estimation

Gabriëlle J. M. De Lannoy *

NASA Goddard Space Flight Center, Code 610.1, Greenbelt Road, Greenbelt, MD 20771, USA

Universities Space Research Association, 7178 Columbia Gateway Drive, Columbia, MD 21046, USA

Rolf H. Reichle

NASA Goddard Space Flight Center, Code 610.1, Greenbelt Road, Greenbelt, MD 20771, USA

* *Corresponding author address:* Gabriëlle J. M. De Lannoy, NASA Goddard Space Flight Center, Code 610.1, Greenbelt Road, Greenbelt, MD 20771, USA.

E-mail: Gabrielle.DeLannoy@nasa.gov, Gabrielle.DeLannoy@ees.kuleuven.be

ABSTRACT

Multi-angle and multi-polarization L-band microwave observations from the Soil Moisture Ocean Salinity (SMOS) mission are assimilated into the Goddard Earth Observing System version 5 (GEOS-5) land surface model, using a spatially distributed ensemble Kalman filter. A variant of this system is also used for the Soil Moisture Active Passive (SMAP) Level 4 soil moisture product. The assimilation involves a forward simulation of brightness temperatures (Tb) for various incidence angles and polarizations, and an inversion of the differences between Tb forecasts and observations into updates to modeled surface and root-zone soil moisture, as well as surface soil temperature. With SMOS Tb assimilation, the unbiased root-mean-square difference between simulations and grid-cell-scale in situ measurements in a few US watersheds during the period of 1 July 2010 - 1 July 2014 is 0.034 m^3/m^3 for both surface and root-zone soil moisture. A validation against grid-cell-scale measurements and point-scale measurements from sparse networks in the US, Australia and Europe demonstrates that the assimilation improves both surface and root-zone soil moisture results over the open loop (no assimilation) estimates in areas with limited vegetation and terrain complexity. At the global scale, the assimilation of SMOS Tb introduces mean absolute increments of 0.004 m^3/m^3 to the profile soil moisture content and 0.7 K to the surface soil temperature. The updates induce changes to energy fluxes and runoff amounting to about 15% of their respective temporal standard deviation.

1. Introduction

Soil moisture, soil temperature and vegetation are important land surface variables in the global weather and climate system. Root-zone soil moisture in particular determines the water availability to plants and the partitioning of water into infiltration, runoff, and evapotranspiration. Estimating root-zone soil moisture at the global scale, however, remains a major challenge.

Global estimates of surface soil moisture can be inferred from satellite-based low frequency passive microwave observations collected by, for example, the current Advanced Microwave Scanning Radiometer 2 (AMSR2) instrument (Imaoka et al. 2010), the Soil Moisture Ocean Salinity (SMOS) mission (Kerr et al. 2010), the Aquarius mission (Le Vine et al. 2007) and the Soil Moisture Active Passive (SMAP) mission (Entekhabi et al. 2010b). However, the utility of spaceborne radiometry is constrained by the limited vertical penetration depth, the coarse spatial resolution, the indirect connection to relevant land surface variables and the intermittent nature of the measurements.

The assimilation of passive microwave measurements, i.e. brightness temperatures (Tb), into land surface models has the potential to add value to these satellite data by (i) increasing the effective vertical penetration depth though propagation of surface information to the root-zone (Galantowicz et al. 1999; Crow and Wood 2003), (ii) increasing the spatial resolution through dynamic downscaling (Reichle et al. 2001a), (iii) increasing the spatial and temporal coverage by interpolation and extrapolation to unobserved times and locations (Reichle and Koster 2003) and (iv) providing enhanced and consistent estimates of various land surface state and flux estimates (Crow and Wood 2003). The objective of this paper is to study these benefits, with exclusion of the spatial downscaling (topic of future research), through the direct and joint assimilation of multi-angle and multi-polarization SMOS Tb into a carefully designed global modeling and assimilation system.

Traditionally, Tb measurements are inverted to soil moisture products prior to use in land surface model applications. The assimilation of soil moisture retrievals in large-scale studies

has been beneficial to improve surface and root-zone soil moisture (Reichle and Koster 2005; Liu et al. 2011; Pan et al. 2012; Draper et al. 2012; De Lannoy et al. 2014c), as well as other land surface variables (Peters-Lidard et al. 2011; Lievens et al. 2015). However, soil moisture retrievals may be inconsistent with model simulations, as they may utilize different land surface parameters and background information including, for example, vegetation, soil texture and surface temperature. Furthermore, errors in retrievals will be correlated to errors in the auxiliary (e.g. temperature) information, which in turn can be expected to be correlated with the background information used in the data assimilation system. Such resulting error cross-covariances are usually ignored in retrieval assimilation.

It is thus natural to consider a direct assimilation of Tb observations into land surface models. Kalman filters have been used to assimilate Tb at various spatial scales, using either synthetic data (Entekhabi et al. 1994; Reichle et al. 2001a; Wilker et al. 2006; Han et al. 2013; Kumar et al. 2015; Carrera et al. 2015) or radiometer data (Galantowicz et al. 1999; Margulis et al. 2002; Crow and Wood 2003; Huang et al. 2013; Loew et al. 2009; Zhang et al. 2011). Variational Tb assimilation has also been a popular technique (Reichle et al. 2001b; Pathmathevan et al. 2003; Balsamo et al. 2006; Jia et al. 2013; Zhao et al. 2013). The crucial component in these Tb assimilation systems is the inclusion of a radiative transfer model (RTM) as the observation operator to connect the soil moisture, soil temperature and vegetation characteristics from the land surface model with Tb predictions.

The current paper explores the joint assimilation of SMOS Tb observations at multiple incidence angles, polarizations and locations at the global scale. While small scale and synthetic studies have been promising, there can be several reasons why the theoretically more attractive direct assimilation of Tb data may not add the expected value for global applications in practice (e.g. Joiner and Dee (2000) for atmospheric applications). The benefit of Tb assimilation in continually improving modeling systems depends on a careful treatment of errors in the assimilation system, which may be complex in large systems that simultaneously assimilate multiple observations.

The modeling and assimilation system used in this paper is the Goddard Earth Observing System version 5 (GEOS-5) land data assimilation system with the Catchment model (CLSM, Koster et al. (2000)) as the land surface model. The GEOS-5 land data assimilation system uses the ensemble Kalman filter (EnKF) and has been used to assimilate a variety of remotely sensed observations separately (Reichle et al. 2014). Recent upgrades to the modeling system include developments in both the land surface model and L-band radiative transfer model (Reichle et al. 2011; De Lannoy et al. 2013, 2014a,b). A variant of this system is used for the operational SMAP Level 4 Surface and Root Zone Soil Moisture (L4_SM) product (Entekhabi et al. 2014) and this paper can be seen as a first assessment of the system for soil moisture estimation through Tb data assimilation.

The main differences between the multi-angular Tb SMOS data assimilation presented in this paper and the SMAP Tb data assimilation performed for the L4_SM product pertain to the nature of the assimilated Tb observations and the difference in spatial resolution of the end products. SMAP data are collected at one incidence angle and with a relatively small instrument error standard deviation ($\sim$1.3 K). In contrast, SMOS provides brightness temperature observations at a range of incidence angles, albeit with a higher instrument error standard deviation ($\sim$4 K). The SMOS soil moisture retrieval algorithm thus uses multi-angular data (Wigneron et al. 2007; Kerr et al. 2012), as does the assimilation algorithm presented here. Moreover, the estimates from the assimilation system in the present paper are at 36 km resolution, whereas the SMAP L4_SM product provides estimates at 9 km resolution.

Section 2 describes the GEOS-5 CLSM and L-band microwave RTM as well as the SMOS observations and ground validation data. Section 3 discusses the data assimilation system, section 4 lists the experiments and validation metrics, and section 5 presents the results.

2. Observations and Model

a. SMOS Brightness Temperature (Tb) Observations

The SMOS mission provides global Tb data at a nominal (3 dB) spatial resolution of 43 km and with global coverage (at either 6:00 am or 6:00 pm local time, i.e. ascending or descending half-orbits, separately) approximately every 3 days. The Microwave Imaging Radiometer with Aperture Synthesis (MIRAS) onboard SMOS is an interferometric sensor and provides multi-angular Tb data at each observed location. Here, we assimilate the multi-angular, horizontally (H) and vertically (V) polarized Tb observations during the period 1 January 2010 through 1 July 2014 and we analyze the results from 1 July 2010 through 1 July 2014. The data are extracted from the MIR_SCLF1C product, with processor version 504 for the years 2010 and 2011, and version 505 from January 2012 onwards.

The various steps involved in the processing of the multi-angular SMOS Tb data are described in De Lannoy et al. (2013) and De Lannoy et al. (2015). The SMOS SCLF1C data are transformed by geometric and Faraday rotation from brightness temperature at the top of the ionosphere and in the antenna reference frame to brightness temperature at the top of the atmosphere and on an Earth-fixed grid. Most importantly, the data are screened extensively using both product-based data quality information and model-based quality control rules (e.g. detection of frozen conditions or heavy precipitation). We limit the Tb data to the exclusively alias-free zone, which leads to relatively narrow swaths and consequently a lower revisit frequency (i.e. global coverage approximately every 6 days for either ascending or descending orbit direction). Data contaminated by radio frequency interference (RFI) are removed when Tb>320 K or guided by product-based flags. Furthermore, the data are spatially mapped onto the 36 km Equal-Area Scalable Earth Grid version 2 (EASEv2, (Brodzik et al. 2014)) and binned per 1^o incidence angle. For example, an observation at 40^o represents the average of all data with incidence angles between 39.5^o and 40.5^o. We assimilate swaths of both H- and V-polarized Tb data at 7 select incidence angles simultaneously: θ=[30^o, 35^o,

40°, 45°, 50°, 55°, 60°]. This range of angles adequately samples the angular signature in the brightness temperature signal. Observations at lower incidence angles are not assimilated because of quality concerns in the data version used here (Martín-Neira et al. 2012).

Regardless of the incidence angle, a circular footprint with a 0.22° radius is assumed, which approximates the area of a 36 km EASEv2 grid cell. This is a simplified approach to the actual field of view which is determined by a spatially variable antenna pattern: about 50% of the signal (3 dB) originates from a 43 km×43 km area, whereas the rest of the signal (with a reduced weight) comes from a larger surrounding area. The weighting by the antenna pattern is included in future SMOS assimilation research and is also implemented for the SMAP L4_SM product.

b. GEOS-5 Land Surface and Radiative Transfer Model

The simulation of Tb involves (i) land surface modeling with the GEOS-5 Catchment land surface model (CLSM, Koster et al. (2000)) and (ii) L-band radiative transfer modeling (RTM) with a tau-omega model (De Lannoy et al. 2013, 2014b). All simulations are performed with a 7.5 minute integration time step, and the computational elements (or tiles) of the modeling system correspond to the 36 km EASEv2 grid cells.

The CLSM version used here is based on the Fortuna 2.5 version of the Modern-Era Retrospective analysis for Research and Applications with improved Land surface variables (MERRA-Land, Reichle et al. (2011)), except here we use the 5 cm surface soil moisture layer depth, the new soil parameterization and the minor code changes of De Lannoy et al. (2014a). The surface layer thus extends from the surface to a depth of 5 cm, the root-zone layer from the surface to a fixed depth of 1 m and the depth of the entire soil profile varies in space between a global minimum of 1.3 m and a maximum of 8.7 m. Surface meteorological forcing data at a 1/2°×2/3° spatial and hourly temporal resolution are taken from MERRA (Rienecker et al. 2011) and bilinearly interpolated to the EASEv2 grid. In some experiments (see later, section 4), the MERRA precipitation is corrected with gage-based precipitation

from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center "Unified" (CPCU) precipitation product (Reichle 2012; Reichle and Liu 2014).

For each computational element, the CLSM model uses three prognostic variables – catchment deficit (catdef), root-zone excess (rzexc) and surface excess (srfexc) – to determine the equilibrium soil moisture profile and deviations from the equilibrium profile in the surface and root-zone. The diagnostic surface soil moisture content (sfmc, 0-5 cm) and root-zone soil moisture content (rzmc, 0-100 cm) are calculated from these three prognostic variables along with the saturated, unsaturated and wilting areas within the tile. In the absence of snow, the surface skin temperature is determined by the area-weighted average temperature across the saturated (tc1), unsaturated (tc2) and wilting (tc4) sub-tile areas. Finally, the ground heat content (ght1) model prognostic variable determines the soil temperature in the uppermost soil layer (tp1). Further details about the model variables can be found in Reichle (2012), with the difference that the surface soil moisture depth is 5 cm in this paper.

The soil moisture, soil temperature, air temperature and climatological vegetation dynamics in the CLSM are used as inputs to the diagnostic zero-order (tau-omega) microwave RTM to simulate L-band Tb. Because RTM parameters strongly impact the climatology of simulated Tb, some key RTM parameters were calibrated by minimizing the bias in the mean and variance between long-term SMOS Tb observations and Tb simulations across a set of incidence angles, polarizations and overpass times (De Lannoy et al. 2013, 2014b). Consequently, the calibrated Tb simulations still show some residual bias at individual incidence angles, polarizations or overpass times, and the bias typically shows a seasonal evolution. Section 3b discusses how these shorter-term biases are addressed inside the assimilation system. Grid cells without sufficient historical SMOS data for RTM calibration are assigned RTM parameter values based on their dominant vegetation class. These nominal parameter values are calculated as global averages across all calibrated pixels per vegetation class, excluding areas where the soil is classified as peat. In this study, the RTM calibration is done over the same 4-year period as the assimilation experiments. The calibration minimizes

climatological biases (De Lannoy et al. 2013), whereas the assimilation addresses random (or short term) errors.

c. Ground Validation Data

The assimilation results are validated during the period 1 July 2010 - 1 July 2014, using independent in situ observations of surface and root-zone soil moisture, as well as surface soil temperature from sparse networks in the US, Australia and Europe, and from intensively monitored validation watersheds (Entekhabi et al. 2014) in the US. All data are subjected to intensive quality control to remove irregularities or sensor trends as discussed in Entekhabi et al. (2014) and De Lannoy et al. (2014a).

The sparse networks include the US Natural Resources Conservation Service (NRCS, Schaefer et al. (2007)) Soil Climate Analysis Network (SCAN), the US Climate Reference Network (USCRN, Diamond et al. (2013); Bell et al. (2013)), the Oznet network in Australia's Murrumbidgee catchment (Smith et al. 2012) and the Soil Moisture Observing System - Meteorological Automatic Network Integrated Application (SMOSMANIA) in France (Albergel et al. 2008; Dorigo et al. 2011). The advantage of SCAN and USCRN is that they offer extensive ground data of soil moisture and temperature in both the surface and root-zone and across a variety of climatological conditions and land surface characteristics. The Oznet and SMOSMANIA sites cover a smaller domain, but complement the validation in different continents.

Surface soil moisture measurements are taken at approximately 5 cm depth. For SCAN and USCRN sites, root-zone soil moisture measurements are a weighted average of measurements at 5 cm, 10 cm, 20 cm and 50 cm depth. For Oznet and SMOSMANIA sites, root-zone measurements are extracted at 45 cm and 30 cm depth, respectively, for lack of sufficient data in other layers. For SCAN and USCRN, sites missing data in any of the root-zone layers are excluded, even if good surface measurements would be available for surface soil moisture validation. In contrast, for the Oznet and SMOSMANIA networks, no cross-masking be-

tween the surface and root-zone validation is included, because of data limitations. Surface soil temperature is validated separately.

Unless noted otherwise, validation sites are limited to 'favorable' areas, i.e. to areas where the relationship between Tb observations and soil moisture is relatively straightforward. Sites are excluded from this category where the maximum climatological Moderate Resolution Imaging Spectroradiometer-based leaf area index (Mahanama et al. 2015) exceeds 5, or if they are within 36 km grid cells that are predominantly covered by forest or shrubland according to the International Geosphere-Biosphere Programme vegetation classification. Finally, sites are excluded in areas with complex topography (determined from GEOS-5 parameters based on HYDRO1k data; Verdin and Greenlee (1996) information), or if the site elevation differs by more than 500 m from the mean elevation of the 36 km grid cell.

The point validation is complemented with a validation using grid-cell-scale in situ data, referred to as data in 'reference grid cells'. In a few United States Department of Agriculture (USDA) watersheds across the US, dense local networks of sensors were installed to calibrate and validate coarse-scale remote sensing observations (Cosh et al. 2008; Jackson et al. 2010). The measurements from these watersheds are part of the core validation sites used to evaluate SMAP data products (Entekhabi et al. 2014). Table 1 lists 10 reference grid cells (36 km) within 7 USDA watersheds. Each reference grid cell has a minimum of 5 individual sensors measuring surface soil moisture for at least 2 years during the validation period. Only 5 of the 10 reference grid cells also have root-zone soil moisture measurements. The lengths of the data records vary but cover most of the 4-year validation period, except for South Fork where the record is limited to 2 years.

3. Data Assimilation

a. Distributed Ensemble Kalman Filter (3D-EnKF)

The EnKF system simultaneously assimilates a set of multi-angular, H- and V-polarized SMOS Tb observations (each with a footprint radius of 0.22^o) located within a circular area with a 1.25^o radius around each 36 km model grid cell. The differences between these multiple Tb observations and their modeled counterparts are used to update the relevant underlying land surface model state variables at each 36 km model grid cell. A schematic of the data assimilation system is shown in Figure 1.

The distributed or 'three-dimensional' (3D) EnKF has been used in earlier land surface data assimilation experiments (Reichle and Koster 2003; De Lannoy et al. 2010; Sahoo et al. 2013). These earlier studies include both spatial smoothing and downscaling of coarse observations to the finer model resolution as part of the 3D filter, whereas in this paper the observation and the model grid resolution are similar (model resolution: EASEv2 36 km, observation footprint radius: 0.22^o) and the 3D filter mainly serves for spatial interpolation and extrapolation. The downscaling to a finer (9 km) resolution to support the SMAP L4_SM product (Entekhabi et al. 2014) will be discussed in future research.

The land surface model $\mathbf{f}(.)$ propagates an ensemble of N initial or analysis state vectors $\hat{\mathbf{x}}^{j+}_{k,i-1}$ from time $i-1$ to an ensemble of N forecast vectors $\hat{\mathbf{x}}^{j-}_{k,i}$ at time i:

$$\hat{\mathbf{x}}^{j-}_{k,i} = \mathbf{f}_{i,i-1}(\hat{\mathbf{x}}^{j+}_{k,i-1}, \mathbf{u}_{k,i-1}, \mathbf{w}^{j}_{k,i-1}), \tag{1}$$

where k denotes space (36 km model grid cell), j denotes an ensemble member ($j = 1, \cdots, N$), $\mathbf{u}_{k,i-1}$ represents the forcings, and $\mathbf{w}^{j}_{k,i-1}$ denotes the model error of (or perturbations to) the j-th ensemble member (section 3c). The ensemble mean model forecast is given by $\hat{\mathbf{x}}^{-}_{k,i} = \frac{1}{N}\Sigma^{N}_{j=1}\hat{\mathbf{x}}^{j-}_{k,i}$. Within the assimilation scheme, the state for a single grid cell k is composed of seven prognostic CLSM variables related to soil moisture and temperature (section 2b), i.e. $\mathbf{x}_{k,i} = [\text{catdef}, \text{srfexc}, \text{rzexc}, \text{tc1}, \text{tc2}, \text{tc4}, \text{ght1}]^{T}_{k,i}$, where T is the vector or matrix transpose. These select variables are expected to be most sensitive to Tb.

In future developments of the system, variables related to vegetation water content could be considered in the state vector.

When SMOS Tb observations $\mathbf{y}_i$ are available at time step i, the state of each ensemble member j is updated as follows:

$$\hat{\mathbf{x}}_{k,i}^{j+} = \hat{\mathbf{x}}_{k,i}^{j-} + \mathbf{K}_{k,i}[\mathbf{y}_i^j - \hat{\mathbf{y}}_i^{j-}]. \tag{2}$$

with $\mathbf{K}_{k,i}$ the Kalman gain, $\hat{\mathbf{y}}_i^{j-} = \mathbf{h}_i(\hat{\mathbf{x}}_i^{j-})$ the ensemble Tb observation predictions, and $\mathbf{h}_i(.)$ the observation operator (see below). The ensemble mean analysis is $\hat{\mathbf{x}}_{k,i}^{+} = \frac{1}{N}\Sigma_{j=1}^{N}\hat{\mathbf{x}}_{k,i}^{j+}$. The vector $\mathbf{y}_i^j$ contains suitably perturbed observations (Burgers et al. 1998) of SMOS H- and V-polarized Tb at multiple incidence angles. These observations are spatially distributed within an influence radius of 1.25° around the model grid cell k (see section 3c). Each of the individual Tb observations ($y_{\lambda,i}^j \in \mathbf{y}_i^j$) has a simulated counterpart $\hat{y}_{\lambda,i}^{j-} \in \hat{\mathbf{y}}_i^{j-}$, with $\hat{y}_{\lambda,i}^{j-} = h_{\lambda,i}(\hat{\mathbf{x}}_i^{j-})$, where $h_{\lambda,i}(.)$ is the observation operator (Reichle et al. 2014), which maps the CLSM state variables to the Tb observation predictions using a radiative transfer model and spatial aggregation, and λ refers to the polarization, incidence angle and location of the individual Tb observations and simulations. The analysis at each 36 km grid cell is thus based on various types (H, V, multiple incidence angles) of spatially distributed Tb simulations within an influence area with a 1.25° radius around the model grid cell, which typically includes several hundred SMOS Tb observations.

Equation 2 inverts the vector of Tb innovations ($\mathbf{y}_i^j - \hat{\mathbf{y}}_i^{j-}$, observations-minus-forecasts residuals) into increments to modeled soil moisture and temperature according to the partitioning given by the Kalman gain $\mathbf{K}_{k,i}$. The Kalman gain utilizes the relative uncertainty of the forecasted prognostic variables (related to soil moisture and soil temperature) and partitions the Tb innovations into the corresponding increments to these prognostic variables. The gain matrix is identical for all ensemble members and determined by:

$$\mathbf{K}_{k,i} = \mathrm{Cov}(\hat{\mathbf{x}}_{k,i}^{-}, \hat{\mathbf{y}}_i^{-}) \left[\mathrm{Cov}(\hat{\mathbf{y}}_i^{-}, \hat{\mathbf{y}}_i^{-}) + \mathbf{R}_i\right]^{-1}, \tag{3}$$

where $\mathrm{Cov}(\hat{\mathbf{x}}_{k,i}^{-}, \hat{\mathbf{y}}_i^{-})$ is the (sample) error covariance (across the ensemble) between the fore-

casted land surface state and the forecasted Tb. Similarly, $\mathrm{Cov}(\hat{\mathbf{y}}_i^-, \hat{\mathbf{y}}_i^-)$ is the (sample) error covariance of the Tb forecasts, and $\mathbf{R}_i$ is the Tb observation error covariance. The Kalman filter only works near optimally with good choices for the error covariances and in the absence of biases. To meet these conditions, biases in observations and simulations are removed prior to assimilation (section 3b) and special attention is given to the determination of the Tb forecast and observation error variances in section 3c and section 3d.

The EnKF system is designed to update root-zone soil moisture in response to observations that are mainly related to surface soil moisture. Thereby, root-zone soil moisture increments rely on ensemble error correlations between the surface and root-zone. Root-zone soil moisture estimates from the analysis are also informed through vertical propagation of surface increments, which depends on the vertical coupling strength in the CLSM (Kumar et al. 2009). The distributed analysis also enables soil moisture and soil temperature updates at unobserved times and locations, so that soil moisture and temperature are updated more frequently than at every overpass. A spatially smooth analysis is ensured through spatially correlated forecast errors, which are suppressed in locations that are farther removed from the analysis grid cell.

b. Biases

The Kalman filter assumes unbiased Tb observations and forecasts, and thus unbiased innovations. Even after RTM calibration, however, there are residual biases between SMOS Tb observations and GEOS-5 Tb forecasts that vary with season, incidence angle, polarization and orbit direction (or overpass time). It is not yet clear how to adequately partition these biases into observation and forecast bias or to come up with appropriate bias models to guide dynamic bias updates (De Lannoy et al. 2007; Reichle et al. 2010; Pauwels et al. 2013; Draper et al. 2014) at the global scale. Therefore, we rely on historical records of the SMOS Tb observations and the corresponding ensemble mean GEOS-5 Tb forecasts to remove seasonally varying climatological bias in the Tb innovations prior to data assimilation.

To this end, we first compute 45-day moving average time series of the Tb observations and ensemble mean forecasts, separately for each 36 km grid cell, incidence angle, polarization, and orbit direction. Next, 4-year averages of these smoothed time series are calculated for each pentad p (5-day period) of the year, yielding a smooth climatological time series (that is, a seasonal cycle) for the Tb observations ($< \mathbf{y} >_p$) and forecasts ($< \hat{\mathbf{y}}^- >_p$). This is done for each grid cell separately, i.e. without spatial smoothing, because neighboring coarse-scale grid cells with dissimilar land surface features would otherwise introduce undesirable bias. Finally, the differences between the Tb observation and forecast climatologies are removed from the Tb innovations for the state update:

$$\hat{\mathbf{x}}_{k,i}^{j+} = \hat{\mathbf{x}}_{k,i}^{j-} + \mathbf{K}_{k,i}[\mathbf{y}_i^j - < \mathbf{y} >_p - (\hat{\mathbf{y}}_i^{j-} - < \hat{\mathbf{y}}^- >_p)]. \quad (4)$$

This approach thus corrects for the first order (i.e. climatological mean) differences between Tb observations and forecasts. We do not correct for seasonal differences in the variability of the observed and simulated Tb anomalies, but instead we assume that the generally higher variability in the SMOS Tb observations can be attributed to observation error (section d).

Figures 2a and b illustrate differences between climatological ensemble mean Tb simulations and SMOS observations at 40° incidence angle on pentad 36 (June 27-July 1) for ascending Tb at H-polarization and descending Tb at V-polarization. The bias is very different for each combination of incidence angle, orbit and polarization, because the RTM was calibrated so that the long-term bias across all assimilated Tb observation types would be minimal. In exceptional areas like the Sahara desert, a known cold model bias is found across all Tb types on average. Figures 2c and d further illustrate that the climatological bias at individual locations also varies in time. Figure 2 is produced using ensemble open loop (model-only, no assimilation) simulations without CPCU precipitation corrections (section 2b). The bias is slightly smaller when CPCU corrections are used.

Note that for soil moisture retrieval assimilation, innovation biases are often removed through cumulative distribution function matching (Reichle et al. 2004), where long-term higher order statistics of observations and simulations are matched, regardless of the time

in the year. In addition to its sensitivity to soil moisture, brightness temperature also depends strongly on the land surface temperature and on vegetation characteristics, both of which typically have a strong seasonal cycle. Bias in brightness temperature therefore varies seasonally, and its correction must also depend on the season. We therefore use a seasonally evolving climatological mean adjustment.

c. Forecast Errors

The forecast error covariances $\mathrm{Cov}(\hat{\mathbf{x}}_{k,i}^{-}, \hat{\mathbf{y}}_{i}^{-})$ and $\mathrm{Cov}(\hat{\mathbf{y}}_{i}^{-}, \hat{\mathbf{y}}_{i}^{-})$ are diagnosed from ensemble model trajectories of the state variables $\hat{\mathbf{x}}_{k,i}^{j-}$ and Tb forecasts $\hat{\mathbf{y}}_{i}^{j-}$, obtained by running the Catchment model with the perturbations listed in Table 2. The precipitation and shortwave radiation are perturbed with multiplicative lognormal perturbations, and longwave radiation receives additive normal perturbations. All forcing perturbations have a temporal correlation length of one day and a spatial correlation length of 0.5°. Moreover, moderate cross-correlations between the perturbed forcings are imposed to ensure physical consistency between the forcing errors.

The prognostic state variables catdef and srfexc (related to soil moisture) are perturbed to mimic errors in model structure and model parameters. Prognostic variables related to surface (skin) and soil temperature are not perturbed explicitly to avoid excessive temperature updates, but a limited uncertainty in temperatures is implicitly created through the perturbation of the radiation and other forcings. Yet, over dry areas with limited vegetation (e.g. Sahara desert), the perturbation in radiation may still result in a large temperature spread.

The perturbations to the model prognostic variables are additive with a temporal correlation length of 3 hours, a spatial correlation length of 0.5° and no cross-correlations. However, weak cross-correlations will develop through balancing during the model simulation; explicitly correlated perturbations to catdef and srfexc were found to excessively update root-zone soil moisture. A new sample of perturbations is calculated every 3 hours and then interpo-

lated and applied at every model time step (7.5 min). During the state updating, spurious long-range forecast error correlations are suppressed beyond 1.25°, i.e. 2.5 times the spatial correlation length, using a Hadamard multiplication of the sample error covariance terms with a distance-dependent and compactly-supported function (Reichle and Koster 2003; Gaspari and Cohn 1999). This localization limits the complexity of the Kalman gain inversion and effectively selects only observations located within a 1.25° radius around the analysis grid cell for assimilation.

Spatially and temporally correlated forcing and state perturbations will cause correlated state forecast errors. The filter adequately accounts for the spatial error correlations by performing a smooth spatial state update. In contrast, despite the temporal correlations, updates are only calculated at single instants, because the land surface model and filtering only rely on the last update to compute the next forecast.

Whereas the perturbation parameters are constant in space and time, the resulting spread in soil moisture and temperature show distinct spatial and temporal patterns, corresponding with the magnitudes of the respective variables. The resulting uncertainty in Tb observation predictions $\hat{\mathbf{y}}_i^{j-} = \mathbf{h}_i(\hat{\mathbf{x}}_i^{j-})$ also exhibits distinct spatial and temporal patterns, with strong inter-angular and inter-polarization cross-correlations, that is, $\mathrm{Cov}(\hat{\mathbf{y}}_i^-, \hat{\mathbf{y}}_i^-)$ has non-negligible off-diagonal elements. The values for the Tb error standard deviations are less than 2 K over forested regions and about 8 K over dry bare soil, such as in the Sahara desert (not shown). This is the expected uncertainty in Tb for soil moisture uncertainties of about 0.01-0.03 $\mathrm{m}^3/\mathrm{m}^3$ (Jackson 1993). The Tb forecast uncertainty over forested areas is small, because Tb is not sensitive to soil moisture perturbations under dense vegetation and soil temperature perturbations are limited by design. Random errors in vegetation (and other land surface and radiative transfer model parameters) are not included in the Tb observation prediction error and are accounted for as representativeness error (part of the observation error, section d). Slowly varying biases in soil moisture, temperature and vegetation are removed as discussed in section b.

d. Observation Errors

In a data assimilation framework, the Tb observations $\mathbf{y}_i$ contain both instrument and representativeness error as part of the total observation error $\mathbf{v}_i$, defined as follows:

$$\mathbf{y}_i = \mathbf{h}_i(\mathbf{x}_i) + \mathbf{v}_i \tag{5}$$

where $\mathbf{x}_i$ refers to the true land surface state variables and $\mathbf{h}_i(.)$ is the imperfect RTM and spatial aggregation. If the observation operator $\mathbf{h}_i(.)$ were perfect, then the observation error $\mathbf{v}_i$ would only consist of instrument errors that have a variance of about 4^2 K^2 for SMOS Tb observations at individual incidence angles (Kerr et al. 2013). Yet, imperfections in the RTM can be attributed to the parameters, structure or auxiliary information (e.g. soil, vegetation). These RTM errors, along with the imperfect match between Tb observations and model predictions in time and space, constitute representativeness errors. The variance of such errors must be added to the instrument error variance to obtain the total 'observation' error variance, i.e. the diagonal of the matrix $\mathbf{R}$. Note that 'observation' error by convention (Eq. 5) thus includes a 'modeling' error component.

The observation error covariance matrix $\mathbf{R}_i$ can be complicated, because multiple observations at different locations and with different incidence angles and polarizations are assimilated simultaneously. For simplicity, we consider a spatially and temporally constant observation error variance of 6^2 K^2 across all angles, polarizations and orbits. This estimate results in near optimal assimilation diagnostics on average across the globe, but not necessarily in individual regions (section 5b1). We further assume an isotropic spatial error correlation length of 0.2^o and a space- and time-invariant exponential inter-angular error cross-correlation function of $0.49\exp(-0.03\ \Delta\theta)$, for angular lags $\Delta\theta \geq 5^o$. The latter function was estimated with a poor-man's approach to adaptive filtering using spatially and temporally averaged statistics of the observations-minus-analyses and observations-minus-forecasts residuals (Desroziers et al. 2005). Partly because of the inter-angular error correlations (and partly because of the smooth dependency of brightness temperature on incidence

angle), assimilating observations at more than the 7 incidence angles used here should not significantly alter the assimilation results. For simplicity, errors are assumed to be uncorrelated between H- and V-polarized observations, even though such correlations can be expected, especially in the representativeness error. In any case, increasing or decreasing error variances can compensate for an under- or overestimation of the error correlations and yield similar results. Note also that the SMOS Level 2 soil moisture retrieval algorithm (Kerr et al. 2013) does not assume any inter-angular observation error correlations in the Tb data.

4. Experiments and Validation

a. Experiments

The experiments consist of two ensemble open loop experiments (model-only, with perturbations) and two corresponding data assimilation experiments (with the same perturbations as the open loop) as listed in Table 3. The first set is performed without and the second set with CPCU precipitation corrections (section 2b). The simulations without precipitation correction are used to study the expected impact of data assimilation in areas without dense precipitation gages, such as in Africa or central Asia. This approach is in line with Liu et al. (2011), where the relative impact of soil moisture assimilation and precipitation corrections was studied. The open loop experiments are denoted as OL and OL_P, without and with precipitation corrections, respectively. The data assimilation experiments are denoted as DA and DA_P, without and with CPCU precipitation corrections, respectively.

b. Validation Metrics

The skill of the assimilation results is assessed using 4-year time series of 3-hour averaged surface and root-zone soil moisture and surface soil temperature extracted at 36 km grid cells that either contain the point sites of the sparse networks or correspond to reference

grid cells (Section 2c). We include all time steps for which in situ observations are available, regardless of whether Tb was assimilated or not at a given time and location. Skill metrics are calculated for the entire period 1 July 2010 - 1 July 2014, and separately for the central 5 months in the warm season of the same 4-year period. For the validation sites in the northern hemisphere, the warm season includes May through September; for the Oznet sites, the warm season includes November through March. For surface soil temperature, the validation is limited to the warm season only and metrics are calculated separately for each 3-hour time interval (0:00-3:00, 3:00-6:00, ..., 21:00-24:00 UTC) before computing an average metric across the eight 3-hour time intervals. This procedure avoids that diurnal effects would otherwise dominate the skill metrics. Furthermore, it will be shown that surface soil temperature updates have only a limited memory and differences in temperature skill are thus best detected when the metric is limited to the warm season, i.e. when Tb observations are actually assimilated (over non-frozen land). Frozen or snow-covered conditions are always excluded.

Skill is measured in terms of time series correlation (R), anomaly time series correlation (anomR), unbiased root-mean-square difference (ubRMSD, Entekhabi et al. (2010a)), and bias. The anomR is calculated as the time series correlation coefficient between the simulations and ground observations after removal of their respective seasonal climatologies. The climatology is calculated for each 3-hour time interval of each day of the year as the 4-year average of the smoothed time series (smoothing window=31 days) of data points at that particular 3-hour time interval. In contrast, the ubRMSD is obtained after removal of the static long-term mean bias between the simulations and the ground observations. The ubRMSD can also be interpreted as the temporal standard deviation in the errors between in situ observations and simulations.

Skill metrics are only calculated if at least 200 data pairs are available for validation. The 95% confidence intervals assume a Student T-distribution for the bias and mean-square difference (MSD), χ^2 for the unbiased mean-square difference (ubMSD) and an asymptotic

normal distribution for R and anomR after a Fisher Z transformation. The confidence intervals take into account the temporal autocorrelation in the 3-hourly time series, which reduces the number of degrees of freedom. To calculate network-averaged skill metrics, the sites within the SCAN and USCRN networks are clustered using a k-means spatial clustering algorithm (MacQueen 1967) based on the similarity in latitude and longitude of individual sites. The clustering approach avoids that densely sampled areas dominate the validation metrics and it also ensures realistic confidence intervals. If each in situ site was treated as an independent validation point, then the average confidence interval CI over all sites would be unrealistically small ($CI = \Sigma_i^N CI_i/N/\sqrt{N}$, with N the total number of sites), and even a minimal change in skill would be wrongly identified as statistically significant. By first clustering neighboring sites that are exposed to similar atmospheric and land surface conditions, we assume that sites in a cluster are correlated so that the confidence interval for a cluster k is calculated as $CI_k = \Sigma_i^{n_k} CI_i/n_k$, with n_k the number of sites within cluster k. Assuming that each cluster adds independent validation information, the network confidence interval then is calculated as $CI = \Sigma_k^K CI_k/K/\sqrt{K}$, with K the number of clusters.

5. Results

a. In Situ Validation

1) Sparse Networks

Figure 3 shows the anomR for the two open loop experiments (OL, OL_P) and the two assimilation experiments (DA, DA_P), averaged across the SCAN and USCRN sites in favorable locations (Section 2), and for the entire period 1 July 2010 - 1 July 2014. For the SCAN surface soil moisture, the average anomR is 0.46 and 0.55 [-] for open loop simulations OL and OL_P, respectively. With assimilation, the anomR values increase to 0.63 and 0.66 [-] for experiments DA and DA_P, respectively. Improvements are also found in the root-zone,

where the anomR is 0.53 and 0.63 [-] for OL and OL_P and the anomR increases to 0.64 and 0.66 [-] for DA and DA_P. For surface soil temperature, the improvements are small by design (section 3c), but statistically significant. At the USCRN sites, the findings are very similar. As expected (Reichle et al. 2011), the use of gage-based precipitation data (OL_P, DA_P) yields soil moisture simulations that are in better agreement with in situ soil moisture observations than simulations which are not informed about in situ precipitation observations (OL, DA).

The improvements in anomR due to Tb assimilation are comparable to those found in earlier studies where retrievals from AMSR-E, ASCAT or SMOS (Liu et al. 2011; Draper et al. 2012; De Lannoy et al. 2014c) were assimilated in earlier versions of the GEOS-5 CLSM. The SMOS Tb assimilation is able to add information that is complementary to the recent improvements in the modeling system (De Lannoy et al. 2014c), even with superior precipitation forcings. Yet, unlike in Liu et al. (2011), here the improvements due to Tb assimilation are smaller when precipitation corrections are used. However, it is difficult to compare the improvements across the various studies, because the metrics and confidence intervals also depend on the experiment period and the temporal resolution of the validation time series.

Figure 4 shows maps for the change in anomR between the open loop simulation and Tb assimilation experiment without precipitation corrections (ΔanomR = anomR(DA) - anomR(OL)) at individual SCAN and USCRN sites in favorable locations as indicated by the green background shading, for the period 1 July 2010 - 1 July 2014. The number of validation sites used here is limited because of the need to estimate the climatology robustly. Moreover, the significance of the changes is limited, especially for root-zone soil moisture, because of the relatively strong autocorrelation in the root-zone soil moisture time series. Despite these limitations, the figure shows that surface soil moisture is improved at almost all sites (92 out of 100) with a cluster-averaged ΔanomR=0.12 [-], while root-zone soil moisture is improved at a majority of sites (70 out of 100) with a cluster-averaged ΔanomR=0.05

[-]. At a few sites, a degradation can be attributed to an erroneous quantification of the vegetation in the RTM or an inadequate horizontal propagation of information.

Table 4 summarizes the skill of soil moisture and temperature estimates for all sparse networks, limited to favorable locations and to only the 4 warm seasons for each network. Tb assimilation is expected to have most impact on soil moisture estimation when the land is not frozen. The metrics are very similar for the SCAN, USCRN, Oznet and SMOSMANIA networks, with generally higher anomR values for experiments with CPCU corrections than those without CPCU corrections. The anomR for the assimilation experiments is always better than the open loop simulations. The improvements are only statistically significant for the SCAN and USCRN networks in terms of surface soil moisture and, in the absence of CPCU precipitation corrections, also for root-zone soil moisture at SCAN sites (ΔanomR=0.12 [-]). Substantial improvements are found for the Oznet and SMOSMANIA networks, but given the limited time period and the limited spatial coverage of these sites, these improvements are not statistically significant.

Especially in terms of ubRMSD, the skill for root-zone soil moisture is always better than that of surface soil moisture (Table 4) because root-zone soil moisture is less variable. Data assimilation reduces the ubRMSD in both the surface and root-zone for all networks, but as already noted above, the assimilation has fewer opportunities to add complementary information when CPCU precipitation corrections are used. The table further shows that the surface soil temperature skill is not altered through Tb data assimilation, because temperature updates only have a limited memory.

The above validation is constrained to areas with moderate topographic complexity and limited vegetation based on coarse-scale model parameters. A detailed investigation of the Tb data assimilation in more complex areas is beyond the scope of this paper. However, it should be mentioned that soil moisture updates are found under relatively dense vegetation, where Tb observations are rather insensitive to soil moisture. The soil moisture updates under vegetation can be explained by a combination of the spatial propagation of information

from less vegetated areas and the relatively high ensemble soil moisture spread in wetter, vegetated, areas (because the larger forecast uncertainty in these areas results in larger increments).

Figure 5 shows the changes in ubRMSD (ΔubRMSD = ubRMSD(DA) - ubRMSD(OL)) for the experiments without precipitation corrections, for all sites with sufficient root-zone data in either favorable or unfavorable areas during the entire period 1 July 2010 - 1 July 2014. The green background again identifies the areas with limited vegetation and topographic complexity, whereas white areas mark complex terrain or (seasonal) dense vegetation.

For the joint SCAN and USCRN networks, 63% of all sites show an improvement in root-zone soil moisture. A solid improvement is found for most sites in the Great Plains (e.g. Oklahoma, Kansas, Nebraska, Dakotas), whereas degradation is found in the Western mountainous areas (e.g. Utah, Nevada) where root-zone soil moisture is difficult to measure and simulate. Surface soil moisture is improved at 77% of the SCAN and USCRN sites, with large improvements in the Great Plains and minor improvements still found in the western half of the US (not shown).

Improvements in the root-zone soil moisture estimates are also found at 72% of the Oznet sites and at 65% of the SMOSMANIA sites (Figure 5). Model soil moisture at most Oznet sites is updated frequently because SMOS observations are abundant in this area (Section 5b). Especially in the Yanco area (largest cluster in Figure 5b), most sensors show improved root-zone soil moisture. Similarly, the ubRMSD in the surface soil moisture is decreased by more than 0.01 m^3/m^3 at the Oznet sites (not shown). In contrast, the impact of the assimilation is very small at the SMOSMANIA sites in both surface (not shown) and root-zone soil moisture, regardless of the location. This is mainly due to the limited number of SMOS Tb data used to update soil moisture because in this area data are often screened out due to frozen conditions, heavy precipitation or RFI. The spatial filter may also propagate information from the nearby Pyrenees mountains into the Piedmont region

where the sensors are located, which may not always be beneficial for the estimation of local conditions.

2) Reference Grid Cells

The skill of the OL, OL_P, DA and DA_P experiments at the reference grid cells is shown in Figure 6 for the anomR, and in Figure 7 for ubRMSD, bias and R. Figure 6 shows that the assimilation generally increases the anomR for surface and root-zone soil moisture, as well as for surface soil temperature for all reference grid cells. The exception is South Fork, where the in situ data record only covers 2 years, artificial drainage installed for agriculture complicates the interpretation of the soil moisture validation results, and insufficient data are available for temperature validation. The average improvement (ΔanomR) for surface soil moisture is 0.18 [-] without precipitation corrections, and 0.08 [-] with precipitation corrections (statistically significant). Without precipitation corrections, these changes are similar to the ΔanomR=0.16 [-] reported in Liu et al. (2011) when assimilating a long time series of AMSR-E soil moisture retrievals into an older version of the GEOS-5 CLSM. Yet, in contrast to Liu et al. (2011), the Tb assimilation in the updated system here yields smaller improvements when superior precipitation forcings are used. Presumably, the reason for this result is the fact that here we use an improved modeling system. Moreover, the length of the time series used here is shorter, making it more difficult to discern statistically significant improvements.

Panels a and d of Figure 7 show that the open loop without precipitation corrections (OL) yields an ubRMSD of 0.042 m^3/m^3 for the surface and 0.039 m^3/m^3 for the root-zone soil moisture across all reference grid cells. With assimilation (DA), the ubRMSD reduces to 0.036 m^3/m^3 for the surface and 0.033 m^3/m^3 for the root-zone, across the available reference grid cells. When including precipitation corrections, the ubRMSD for the open loop (OL_P) is 0.038 m^3/m^3 for the surface and 0.033 m^3/m^3 for the root-zone soil moisture across all reference grid cells. With assimilation (DA_P), the ubRMSD reduces to 0.034 m^3/m^3 for the

surface and increases slightly to 0.034 m^3/m^3 for the root-zone.

The assimilation slightly decreases the ubRMSD in surface soil temperature (not shown) at all reference grid cells except South Fork, which is a combined effect of directly updating the soil temperature and changes in fluxes resulting from soil moisture updates. On average, the ubRMSD for the open loop and data assimilation experiments is around 2 K, with higher values (3.5 K) in Walnut Gulch and lower values (1.2 K) in Little River (not shown).

Figure 7b and e show the bias values for surface and root-zone soil moisture. The local bias is larger than the ubRMSD, which is expected, when considering that the global land surface model relies on ancillary information for texture and vegetation. With CPCU precipitation corrections, the bias is slightly larger than without such corrections. This was also found for SCAN and USCRN sites (not shown). By design, the assimilation changes the bias only minimally (except in South Fork, where the validation period is too short) and these small changes can be attributed to the spatial nature of the filter. The Tb assimilation further introduces a significant improvement in R for the surface soil moisture (Figure 7c). For the root-zone, the improvement in R due to Tb assimilation alone is not statistically significant (Figure 7f), because of the limited time period, but the improvement is as high as 0.23 [-] when no CPCU precipitation corrections are used.

As an example, Figure 8 shows soil moisture time series for experiments OL_P and DA_P along with in situ observations at Fort Cobb for each July and August in the 4-year experiment period. Due to Tb assimilation, the responses to the precipitation events of July 2010 and August 2013 are better simulated in the root-zone. In August 2010, the SMOS observations are anomalously warm relative to a 4-year climatology, suggesting drier than usual conditions. The assimilation thus changes surface and root zone soil moisture towards the in situ observations. Some irregularities in the surface soil moisture indicate that some updates may be too strong, because the observation error standard deviation of 6 K may be too low at this location.

b. Large Scale Results

In this section, global assimilation diagnostics at the 36 km model resolution are presented to evaluate the quality of the multi-angular SMOS Tb assimilation system. Only simulations with CPCU precipitation corrections are presented for simplicity.

1) Innovations (Observation Space)

Figure 9a shows a map of the average number of ascending or descending half-orbits per day for which SMOS Tb observations of at least one polarization and incidence angle were assimilated. On average, there is one set (multi-angle, 2 polarizations) of Tb observations every 3 days (or every 6 days per orbit direction, see also section 2a), with fewer data available in mountainous and northern areas with prolonged freezing conditions. Because of RFI contamination, data are missing in large portions of Asia, Eastern Europe and the Middle East.

The temporal mean in the bias-corrected (section 3b) Tb innovations is close to 0 K by design (not shown). The temporal standard deviations in the normalized Tb innovations $[\mathbf{y}_i - \hat{\mathbf{y}}_i^-]_\lambda / \sqrt{[\mathbf{R}_i + \mathrm{Cov}(\hat{\mathbf{y}}_i^-, \hat{\mathbf{y}}_i^-)]_{\lambda\lambda}}$, averaged across all angles and polarizations, is shown in Figure 9b for experiment DA_P. This metric is used to verify a necessary condition for the optimality of the filter operation (Reichle et al. 2002): values close to 1 indicate reasonable choices for forecast and observation error variances, whereas values less than 1 indicate overestimated Tb observation or forecast errors (or both), and values larger than 1 suggest underestimated errors.

Figure 9b shows that experiment DA_P leads to areas where the metric is much smaller or larger than 1. For example, in the central agricultural area of the US, the standard deviation of the normalized innovations is larger than 1, which indicates that the DA system underestimates Tb observation and/or forecast errors. The calibration of the RTM (De Lannoy et al. 2014c) revealed higher errors in Tb predictions over cropland and consequently,

it can be expected that $\sqrt{[\mathbf{R}]_{\lambda\lambda}}$=6 K underestimates the representativeness error as part of the observation error. Over the Sahara desert, Figure 9b indicates a large overestimation of Tb observation and/or forecast errors. Here, the brightness temperature forecast errors are most likely overestimated because of the large ensemble spread in surface soil temperature over dry areas with little vegetation. The diagnostic metric shown in Figure 9b can be improved by tuning either the forecast or observation errors alone, as is typically done in adaptive filtering techniques. However, we found that this does not suffice to realistically balance both forecast and observation errors in every region of the globe. Further research is needed to optimize the global system.

2) Increments (State Space)

Figure 10a shows the average number of increments per day applied at each location during the 4 validation years. Increments are counted when a value larger than 10^{-5} mm is found for any soil moisture component (catdef, srfexc, rzexc), or a value larger than 10^{-3} K for surface soil temperature. In observed areas, increments are applied every 1-2 days, which is more frequently than the number of overpasses per day (Figure 9a). The 3D-EnKF effectively applies increments outside the observed swaths (up to a spatial error correlation localization distance of 1.25^o) and in areas where observations are not available or screened out after quality control. It should be noted that the number of increments varies seasonally, with most increments during the summer and a reduced number of increments in the winter.

Figure 10b shows the temporal standard deviation in the increments to the total soil water column (prmc), expressed in volumetric fractions (i.e. accounting for variable profile depths), for the assimilation experiment DA_P. Similarly, Figure 10c shows the standard deviation in the surface soil temperature (tp1) increments. Not shown are the mean increments, which are negligible, because the system is nearly unbiased by design. The standard deviation in the increments is small when spatially and temporally averaged, because many very small increments are associated with distant observations (3D-EnKF). This is especially

true in coastal areas where updates are solely related to distant observations (because Tb observations near water are removed during quality control). For experiment DA_P, the temporal standard deviation in prmc increments amounts to 3.8 10^{-3} m^3/m^3 for the globe (Figure 10b), with some areas showing distinctly larger or smaller absolute increments. The larger increments are often associated with underestimated observation errors (Figure 9b). The global average of the surface soil temperature increments is 0.65 K, with higher values in drier areas with limited vegetation and Tb forecast uncertainties that are perhaps too large due to excessive temperature uncertainties. The temporal standard deviation in the temperature and moisture increments is larger in the summer and reduced in the winter (not shown).

3) Other Variables

Next, we analyze the impact of Tb data assimilation on other land surface variables, including the latent and sensible heat fluxes and runoff, across the globe (Figure 11) and for North America (20^o-50^o N, Table 5). Figure 11 measures impact in terms of the RMSD between the open loop (OL_P) and the assimilation integration (DA_P) over the 4-year time series. For North America only, Table 5 provides the long-term mean and standard deviation values of the open loop OL_P simulation and differences from the corresponding DA_P values separately for each season and for the entire year. Table 5 also provides the ubRMSD values between OL_P and DA_P. Table 5 shows very small differences (bias) in (absolute) long-term means and standard deviations for most variables, which implies that the RMSD is close to the ubRMSD.

The globally averaged RMSD values between OL_P and DA_P are very small (Figure 11), because some regions do not receive any updates for lack of Tb observations of sufficient quality (section 2a). For surface and root-zone soil moisture, the values of the global RMSD are 0.02 m^3/m^3 and 0.01 m^3/m^3, respectively (Figure 11a,b). The RMSD is higher in northern latitudes where the CPCU precipitation corrections are tapered (no corrections

above 62.5^o N) and data assimilation has more impact. When focusing on North America, which is well observed by SMOS, Table 5 reveals that the ubRMSD per season amounts to less than 10% of the open loop mean soil moisture, but to around 50% of the open loop standard deviation. As expected, the ubRMSD in surface soil temperature is very small and only amounts to about 5% of the open loop standard deviation.

The RMSD between OL_P and DA_P in the energy fluxes has values of about 9-10 W/m^2 across the globe for sensible and latent heat, with more pronounced effects in the central US and the Southern Hemisphere. Table 5 indicates that the ubRMSD in energy fluxes is about 20% of the mean open loop values and about 10-20% of the mean temporal standard deviation, with larger changes in the summer and fall months.

The RMSD between OL_P and DA_P in total runoff is about 0.27 mm/day across the globe (Figure 11f) with large RMSD values in the wetter eastern part of North America, where most of the soil moisture updates are applied. The ubRMSD values for North America (Table 5) reach up to 80% of the mean seasonal values and reach up to 20% of the temporal standard deviation in the fall months. Figure 11f shows very high RMSD values over the Amazon and the Indonesia Archipelago where the model produces very high runoff peaks. Small updates in soil moisture over these areas result in large absolute changes in runoff, but the changes are small relative to the magnitude of the runoff.

4) Uncertainty

One of the advantages of the EnKF is its ability to provide ensemble error estimates for simulated variables. Figures 12a, c and e show the analysis error standard deviation in soil moisture and surface soil temperature for the assimilation integration DA_P (enstd^a), averaged over all 3-hourly timesteps during 1 July 2010 - 1 July 2014. The estimated uncertainty in surface soil moisture is about 0.02 m^3/m^3 across the globe and higher in areas with limited or no assimilation updates (e.g. Asia, northern latitudes). The uncertainty is also high, for example, in the dry western US, where the assimilation resulted in smaller

absolute soil moisture increments (Figure 10b). The uncertainty in root-zone soil moisture is smaller (about 0.01 m^3/m^3 across the globe), because the uncertainties in the forcings are dampened. For soil temperature, the ensemble uncertainty is mostly less than 1 K and limited by design (section 3c), but high temperature uncertainties are found in dry areas with limited vegetation. Some coastal grid cells show a high uncertainty, because the assimilation only affects these locations through small updates based on distant Tb observations.

Figure 12 also verifies the expectation that data assimilation reduces the uncertainty of the estimated land surface variables. Specifically, Figures 12b, d and f show the ratio of the temporal mean ensemble standard deviation for the assimilation integration DA_P (enstd^a) to the ensemble standard deviation in the model-only open loop simulation OL_P (enstd^m). Again, the ensemble standard deviations are averaged across all 3-hourly timesteps during 1 July 2010 - 1 July 2014. For surface soil moisture, the averaged ratio is 0.85 across the globe. In some wetter and well observed areas, the estimated uncertainty is halved by the assimilation (e.g. parts of the central US), whereas the uncertainty is only marginally changed in boreal regions and unaltered across much of Asia and south-east Europe where no observations are available for assimilation. In some areas with dense vegetation and where Tb is only marginally sensitive to soil moisture, e.g. in central Africa, the uncertainty is only marginally changed. The root-zone soil moisture shows a similar pattern, but with a greater reduction in uncertainty, i.e. a ratio of 0.74 across the globe. The reduction is greater than in surface soil moisture, because root-zone soil moisture has a longer memory and is less directly exposed to perturbations in the meteorological forcings. In terms of surface soil temperature, only a minor contraction of the ensemble spread is noticed when averaged across all 3-hourly timesteps, because of the limited memory of (surface) soil temperature updates in the modeling system.

6. Conclusions

The direct assimilation of multi-angle and multi-polarization SMOS Tb observations into the GEOS-5 land surface model is evaluated for global soil moisture estimation. The assimilation uses a distributed ensemble Kalman filter with a temporally variable Tb bias mitigation, a system that is also used for the SMAP L4_SM product. The distributed aspect of the filter introduces updates that are spatially extrapolated into unobserved areas. The multi-dimensional aspect of the observations is a unique feature of this paper: for each location on Earth, several hundreds of SMOS Tb observations may be assimilated simultaneously.

Two types of assimilation experiments are performed, one using MERRA forcings and one using MERRA forcings corrected with CPCU gage precipitation. In both cases, the multi-angular SMOS Tb assimilation yields better results than an open loop simulation in terms of the anomaly time series correlation (anomR) and unbiased root-mean-square error (ubRMSD). Improvements are found in both the surface and root-zone soil moisture for sites in sparse networks (SCAN, USCRN, SMOSMANIA, Oznet) and reference grid cells across the US, limited to areas with mild topography and limited vegetation. Tb data assimilation thus effectively helps to propagate surface information to the root-zone. The magnitudes of the improvements in anomR due to SMOS Tb assimilation are comparable to what is reported in studies assimilating longer time series of soil moisture retrievals from AMSR-E or ASCAT in older versions of the GEOS-5 CLSM (Liu et al. 2011; Draper et al. 2012). Yet, an exact comparison is not possible, because the statistics depend on the experiment period and temporal resolution of the validation time series.

With SMOS Tb data assimilation and when using gage-corrected (CPCU) precipitation, the ubRMSD between simulations and grid-cell-scale in situ measurements (also used in the core validation of SMAP products) is 0.034 m^3/m^3 for surface soil moisture at 10 reference grid cells across the US, and also 0.034 m^3/m^3 for root-zone soil moisture at the 5 reference grid cells with such data. Because the benefit of Tb assimilation is found to be larger when

no precipitation corrections are used, we expect that SMOS (or SMAP) assimilation has a greater impact on soil moisture estimates over regions such as Africa and Asia, where both high-quality meteorological information (including precipitation gage measurements) and in situ validation data are unavailable.

The globally averaged values of the time series standard deviation in the increments for the total soil water column and for surface soil temperature are 0.004 m^3/m^3 and 0.7 K, respectively. The values are very small when averaged in time and space, because many small increments are associated with distant updates in the 3D filter (extrapolation). Yet, the corresponding mean absolute difference in soil moisture between the open loop and assimilation estimates can reach up to 50% of the temporal variability, and the energy and runoff fluxes are changed by roughly 15% relative to the temporal variability per season.

The EnKF provides global uncertainty estimates for all variables in the land surface data assimilation system. With assimilation, the globally averaged uncertainty for the analyzed surface and root-zone soil moisture is 0.02 m^3/m^3 and 0.01 m^3/m^3, respectively, which is a fraction of the uncertainty without any assimilation (0.85 for global surface soil moisture and 0.74 for global root-zone soil moisture).

To summarize, SMOS Tb observations are predominantly related to surface soil moisture and temperature and vegetation conditions. The Tb data assimilation successfully (i) increases the effective vertical penetration depth through propagation of surface information to the root-zone, (ii) increases the spatial and temporal coverage by interpolation and extrapolation to unobserved times and locations, and (iii) provides consistent estimates of various land surface state and flux estimates with reduced uncertainty. In future research and in the forthcoming SMAP L4_SM product, the spatial resolution of the land model within the assimilation system and therefore the resolution of the assimilation estimates will be increased to 9 km. The analysis component of this future system will thus downscale the coarser (36 km) brightness temperature observations to the finer (9 km) resolution of the modeling system.

Acknowledgments.

The authors thank Yann Kerr for his assistance with the SMOS data, Mike Cosh and Tom Jackson for providing the in situ data for the SMAP core validation watersheds, Jeffrey Walker and Xiaoling Wu for providing a part of the Oznet data, Qing Liu for helping with the quality control of sparse network data, and the reviewers and Hans Lievens for suggestions to edit the paper. The study was supported by the NASA Soil Moisture Active Passive (SMAP) mission. Special thanks go to our SMAP L4_SM collaborators Randy Koster, Wade Crow, John Kimball, and Qing Liu. Computational resources were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at the Goddard Space Flight Center.

REFERENCES

Albergel, C., et al., 2008: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in situ observations and model simulations. *HESS*, **12**, 1323–1337.

Balsamo, G., F. M. J, S. Belair, and G. Deblonde, 2006: A global root-zone soil moisture analysis using simulated L-band brightness temperature in preparation for the hydros satellite mission. *Journal of Hydrometeorology*, **7**, 1126–1146.

Bell, J., et al., 2013: U.S. climate reference network soil moisture and temperature observations. *Journal of Hydrometeorology*, **14**, 977–988.

Brodzik, M. J., B. Billingsley, T. Haran, B. Raup, and M. Savoie, 2014: Correction: Incremental but significant improvements for earth-gridded data sets. *ISPRS International Journal of Geo-Information*, **3**, 1154–1156.

820 Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: On the analysis scheme in the
821 ensemble Kalman filter. *Monthly Weather Review*, **126**, 1719–1728.

822 Carrera, M., S. Belair, and B. Bilodeau, 2015: The Canadian land data assimilation system
823 (CALDAS): Description and synthetic evaluation study. *Journal of Applied Meteorology*,
824 **16**, 1293–1314.

825 Cosh, M. H., T. J. Jackson, S. Moran, and R. Bindlish, 2008: Temporal persistence and
826 stability of surface soil moisture in a semi-arid watershed. *Remote Sensing of Environment*,
827 **112**, 304–313.

828 Crow, W. T. and E. F. Wood, 2003: The assimilation of remotely sensed soil brightness
829 temperature imagery into a land surface model using ensemble Kalman filtering: A case
830 study based on ESTAR measurements during SGP97. *Advances in Water Resources*, **26**,
831 137–149.

832 De Lannoy, G., R. Koster, R. Reichle, S. Mahanama, and Q. Liu, 2014a: An updated treat-
833 ment of soil texture and associated hydraulic properties in a global land modeling system.
834 *Journal of Advances in Modeling Earth Systems*, **6**, 23p, doi:10.1002/2014MS000330.

835 De Lannoy, G., R. Reichle, P. Houser, V. Pauwels, and N. Verhoest, 2007: Correcting
836 for forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water*
837 *Resources Research*, **43 (W09410)**, 1–14, doi:10.1029/2006WR00544.

838 De Lannoy, G., R. Reichle, and V. Pauwels, 2013: Global calibration of the GEOS-5 L-
839 band microwave radiative transfer model over non-frozen land using SMOS observations.
840 *Journal of Hydrometeorology*, **14**, 765–785, http://dx.doi.org/10.1175/JHM-D-12-092.1.

841 De Lannoy, G., R. Reichle, J. Peng, Y. kerr, R. Castro, E. Kim, and Q. Liu, 2015: Converting
842 between SMOS and SMAP level-1 brightness temperature observations over nonfrozen
843 land. *IEEE Geosceinece and Remote Sensing Letters*, **12**, 1908–1912.

De Lannoy, G., R. Reichle, and J. Vrugt, 2014b: Uncertainty quantification of GEOS-5 L-band radiative transfer model parameters using Bayesian inference and SMOS observations. *Remote Sensing of Environment*, **148**, 146–157, doi:10.1016/j.rse.2014.03.030.

De Lannoy, G. J. M., P. de Rosnay, and R. H. Reichle, 2014c: Soil moisture data assimilation. *Handbook of Hydrometeorological Ensemble Forecasting*, Q. Duan, F. Pappenberger, J. Thielen, A. Wood, H. Cloke, and J. Schaake, Eds., Springer Verlag, XXX, in press.

De Lannoy, G. J. M., R. H. Reichle, P. R. Houser, K. R. Arsenault, N. E. C. Verhoest, and V. R. N. Pauwels, 2010: Satellite-scale snow water equivalent assimilation into a high-resolution land surface model. *Journal of Hydrometeorology*, **11 (2)**, 352–369, doi:10.1175/2009JHM1192.1.

Desroziers, G., L. Berra, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Q. Journal of the Royal Meteorological Society*, **131**, 3385–3396.

Diamond, H., et al., 2013: U.S. climate reference network after one decade of operations: status and assessment. *BAMS*, **94 (4)**, 485–498.

Dorigo, W. A., et al., 2011: The international soil moisture network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, **15 (5)**, 1675–1698.

Draper, C., R. Reichle, G. D. Lannoy, and B. Scarino, 2014: A dynamic approach to addressing observation-minus-forecast mean bias in a land surface skin temperature data assimilation system. *Journal of Hydrometeorology*, **X**, X.

Draper, C. S., R. H. Reichle, G. J. M. D. Lannoy, and Q. Liu, 2012: Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters*, **39**, L04 401.

867 Entekhabi, D., J. Galantowicz, and E. Njoku, 1994: Solving the inverse problem for soil
868 moisture and temperature profiles by sequential assimilation of multifrequency remotely
869 sensed observations. *IEEE Transactions on Geoscience and Remote Sensing*, **32 (2)**, 438–
870 448.

871 Entekhabi, D., R. H. Reichle, R. D. Koster, and W. T. Crow, 2010a: Performance metrics
872 for soil moisture retrievals and application requirements. *Journal of Hydrometeorology*, **11**,
873 832–840.

874 Entekhabi, D., S. Yueh, P. O'Neill, and K. Kellogg, 2014: *SMAP Handbook*. JPL 400-1567,
875 Pasadena, CA, USA.

876 Entekhabi, D., et al., 2010b: The Soil Moisture Active and Passive (SMAP) mission. *Pro-*
877 *ceedings of the IEEE*, **98 (5)**, 704–716.

878 Galantowicz, J., D. Entekhabi, and E. G. Njoku, 1999: Tests of sequential data assimilation
879 for retrieving profile soil moisture and temperature from observed L-band radiobrightness.
880 *IEEE Transactions on Geoscience and Remote Sensing*, **37 (4)**, 1860–1870.

881 Gaspari, G. and S. Cohn, 1999: Construction of correlation functions in two and three
882 dimensions. *Quarterly Journal of the Royal Meteorological Society*, **125**, 723–757.

883 Han, X., H.-J. H. Franssen, X. Li, Y. Zhang, C. Montzk, and H. Vereecken, 2013: Joint
884 assimilation of surface temperature and L-band microwave brightness temperature in land
885 data assimilation. *Vadose Zone Journal*, **12 (3)**.

886 Huang, C., X. Li, L. Lu, and J. Gu, 2013: Experiments of one-dimensional soil moisture
887 assimilation system based on ensemble Kalman filter. *Remote Sensing of Environment*,
888 **112**, 888–900.

889 Imaoka, K., M. Kachi, M. Kasahara, N. Ito, K. Nakagawa, and T. Oki, 2010: Instrument

performance and calibration of AMSR-E and AMSR2. *International Archives of the Photogrammetry, Remote Sensing and Special Information Science*, **38**, 8.

Jackson, T., 1993: Measuring surface soil moisture using passive microwave remote sensing. *Hydrological Processes*, **7 (2)**, 139–152.

Jackson, T. J., et al., 2010: Validation of advances microwave scanning radiometer soil moisture. *IEEE Transactions on Geoscience and Remote Sensing*, **48 (2)**, 4256–4272.

Jia, B., X. Tian, Z. Xie, J. Liu, and C. Shi, 2013: Assimilation of microwave brightness temperature in a land data assimilation system with multi-observation operators. *Journal of Geophysical Research*, **118**, 1–14.

Joiner, J. and D. P. Dee, 2000: An error analysis of radiance and suboptimal retrieval assimilation. *Q.J.R. Meteorological Soc.*, **126**, 1495–1514.

Kerr, Y., P. Waldteufel, P. Richaume, P. Ferrazzoli, and J.-P. Wigneron, 2013: SMOS level 2 processor for soil moisture algorithm theoretical basis document (ATBD) v1.3h. Tech. rep., Toulouse, SM-ESL (CBSA), 141 pp. Tech Note SO-TN-ESL-SM-GS-0001.

Kerr, Y., et al., 2010: The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE*, **98 (5)**, 666–687.

Kerr, Y., et al., 2012: The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, **50 (5)**, 1384–1403.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model 1. model structure. *Journal of Geophysical Research*, **105 (D20)**, 24 809–24 822.

Kumar, S., K. Harrison, C. D. Peters-Lidard, J. Santanello, and D. Kirschbaum, 2015: Assessing the impact of L-band observations on drought and flood risk estimation: A

decision-theoretic approach in an OSSE environment. *Journal of Hydrometeorology*, **15**, 2140–2156.

Kumar, S., R. H. Reichle, R. D. Koster, W. T. Crow, and C. D. Peters-Lidard, 2009: Role of subsurface physics in the assimilation of surface soil moisture observations. *Journal of Hydrometeorology*, **10**, 1534–1547, doi:10.1109/MC.2008.511.

Le Vine, D., G. S. E. Lagerloef, F. R. Colomb, S. H. Yueh, and F. A. Pellerano, 2007: An instrument to monitor sea surface salinity from space. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 2040–2050.

Lievens, H., et al., 2015: SMOS soil moisture assimilation for improved stream flow simulation in the Murray Darling basin, Australia. *Remote Sensing of Environment*, **168**, 146–162.

Liu, Q., R. H. Reichle, R. Bindlish, M. H. Cosh, W. T. Crow, R. de Jeu, G. J. M. D. G. J. Huffman, and T. J. Jackson, 2011: The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system. *Journal of Hydrometeorology*, **12**, 750–765.

Loew, A., M. Schwank, and F. Schlenz, 2009: Assimilation of an L-band microwave soil moisture proxy to compensate for uncertainties in precipitation data. *TGRS*, **47**, 2606 –2616.

MacQueen, J., 1967: Some methods for classification and analysis of multivariate observations. University of California Press, 281–297.

Mahanama, S. P., et al., 2015: Land boundary conditions for the goddard earth observing system model version 5 (geos-5) climate modeling system - recent updates and data file descriptions. Tech. rep., National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA, 39 pp.

937 Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne, 2002: Land data assimi-
938 lation of soil moisture using measurements from the Southern Great Plains 1997 Field
939 Experiment. *Water Resources Research*, **38 (12)**, 35.1–35.18.

940 Martín-Neira, M., et al., 2012: SMOS instrument performance and calibration. *Proceedings*
941 *of IGARSSS*, München, Germany.

942 Pan, M., A. Sahoo, W. Wood, A. Al Bitar, D. Leroux, and Y. Kerr, 2012: An initial assess-
943 ment of SMOS derived soil moisture over the continental United States. *Journal of Selected*
944 *Topics in Earth Observations and Remote Sensing*, **99**, 10.1109/JSTARS.2012.2194477.

945 Pathmathevan, M., T. Koike, X. Lin, and H. Fujii, 2003: A simplified land data assimila-
946 tion scheme and its application to soil moisture experiments in 2002 (SMEX02). *Water*
947 *Resources research*, **39 (12)**, SWC6.1–6.20.

948 Pauwels, V., G. De Lannoy, H.-J. Hendricks Franssen, and H. Vereecken, 2013: Simultaneous
949 estimation of model state variables and observation and forecast biases using a two-stage
950 hybrid Kalman filter. *HESS*, **10**, 5169–5224.

951 Peters-Lidard, C., S. V. Kumar, D. M. Mocko, and Y. Tian, 2011: Estimating evapotran-
952 spiration with land data assimilation systems. *Hydrological Processes*, **25**, 3979–3992.

953 Reichle, R., G. De Lannoy, B. Forman, C. Draper, and Q. Liu, 2014: Connecting satellite
954 observations with water cycle variables through land data assimilation: Examples using
955 the NASA GEOS-5 LDAS. *Surveys of Geophysics*, **35**, 577–606.

956 Reichle, R., S. V. Kumar, S. P. P. Mahanama, R. D. Koster, and Q. Liu, 2010: Assimilation
957 of satellite-derived skin temperature observations into land surface models. *Journal of*
958 *Hydrometeorology*, **11**, 1103–1122, doi:10.1175/2010JHM1262.1.

959 Reichle, R. H., 2012: The MERRA-Land data product (version 1.2), GMAO office note

no. 3. Tech. rep., NASA Global Modeling and Assimilation Office. Available at `http://gmao.gsfc.nasa.gov/pubs/office_notes/`.

Reichle, R. H., D. Entekhabi, and D. B. McLaughlin, 2001a: Downscaling of radio brightness measurements for soil moisture estimation: A four dimensional variational data assimilation approach. *Water Resources Research*, **37 (9)**, 2353–2364.

Reichle, R. H. and R. Koster, 2003: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *Journal of Hydrometeorology*, **4**, 1229–1242.

Reichle, R. H. and R. Koster, 2005: Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model. *Geophysical Research Letters*, **32**, L0204.1–L0204.4.

Reichle, R. H., R. Koster, J. Dong, and A. Berg, 2004: Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation. *Journal of Hydrometeorology*, **5**, 430–442.

Reichle, R. H., R. D. Koster, G. J. M. De Lannoy, B. A. Forman, Q. Liu, S. P. P. Mahanama, and A. Toure, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *Journal of Climate*, **24**, 6322–6338.

Reichle, R. H. and Q. Liu, 2014: Observation-corrected precipitation estimates in geos-5. Tech. rep., National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA, 18 pp.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2001b: Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Transactions on Geoscience and Remote Sensing*, **39 (8)**, 1708–1718.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002: Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, **130 (1)**, 103–114.

Rienecker, M. M., et al., 2011: MERRA - NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, **24 (14)**, 3624–3648, doi:10.1175/JCLI-D-11-00015.

Sahoo, A., G. J. M. De Lannoy, R. H. Reichle, and P. R. Houser, 2013: Assimilation and downscaling of satellite observed soil moisture over the Little River experimental watershed in Georgia, USA. *Advances in Water Resources*, **52**, 19–33.

Schaefer, G. L., M. H. Cosh, and T. J. Jackson, 2007: The USDA natural resources conservation service soil climate analaysis network (scan). *Journal of Atmospheric and Oceanic Technology*, **24**, 2073–2077.

Smith, A., et al., 2012: The Murrumbidgee soil moisture monitoring network data set. . *Water Resources Research*, **48**, W07 701.1–6.

Verdin, K. and S. Greenlee, 1996: Development of continental scale digital elevation models and extraction of hydrographic features. in: Proceedings, third international conference/workshop on integrating gis and environmental modeling. National Center for Geographic Information and Analysis, Santa Fe, New Mexico.

Wigneron, J., et al., 2007: L-band microwave emission of the biosphere (L-MEB) model: Description and calibration against experimental data sets over crop fields. *Remote Sensing of Environment*, **107**, 639–655.

Wilker, H., M. Drusch, G. Seuffert, and C. Simmer, 2006: Effects of the near-surface soil moisture profile on the assimilation of L-band microwave brightness temperature. *Journal of Hydrometeorology*, **7**, 433–442.

Zhang, S., J. Shi, and Y. Dou, 2011: A soil moisture assimilation scheme based on the microwave land emissivity model and the community land model. *International Journal of Remote Sensing*, **33 (9)**, 2770–2797.

1008 Zhao, L., K. Yang, J. Qin, and Y. Chen, 2013: Optimal exploitation of amsr-e signals for
1009 improving soil moisture estimation through land data assimilation. *IEEE Transactions on*
1010 *Geoscience and Remote Sensing*, **51 (1)**, 399–410.

List of Tables

Table 5 Spatial average of the long-term and seasonal statistics for experiment OL_P and differences with experiment DA_P for various land surface variables. OL_P statistics include the temporal mean (mean OL_P) and standard deviation (stdv OL_P). Difference statistics include the difference in the mean, mean of the absolute differences, difference in the standard deviation (MDiff, MabsDiff, SDiff; DA_P minus OL_P) and the ubRMSD between DA_P and OL_P results. Statistics are calculated as the spatial average of the metrics for North America (20-50^o N) during the period 1 July 2010 - 1 July 2014. 48

TABLE 1. 36 km reference grid cells within SMAP core validation watersheds (Entekhabi et al. 2014). Each reference grid cell contains a minimum of 5 surface soil moisture sensors and a maximum of N sensors. The latitude and longitude refer to the center of the EASEv2 36 km grid cells. Sites marked with an asterisk also provide measurements in the root-zone layer.

Watershed	State (US)	Reference grid cell	Latitude (° N)	Longitude (° W)	Maximum N sensors
Reynolds Creek	ID	RC1	43.33	116.70	6
		RC2	42.95	116.70	9
Walnut Gulch	AZ	WG1	31.96	110.73	6
		WG2	31.62	110.35	14
		WG3	31.62	109.98	22
Little Washita	OK	LW*	34.99	98.03	15
Fort Cobb	OK	FC*	35.34	98.40	10
Little River	GA	LR1*	31.62	83.84	15
		LR2*	31.62	83.46	12
South Fork	IA	SF*	42.57	93.55	12

TABLE 2. Ensemble perturbations to forcing and model prognostic variables. The perturbation type is either additive (A) or multiplicative (M), with the standard deviation given for a normal or lognormal distribution, respectively. The time series correlation (temporal corr) is applied to a first order autoregressive model. The spatial correlation scale (spatial corr) is isotropic. Perturbations to prognostic variables are uncorrelated with each other and with the forcing perturbations.

	type	standard deviation	temporal corr.	spatial corr.	cross-corr. with perturbations in		
					P	SW	LW
Precipitation (P)	M	0.5 [-]	24 h	0.5°	n/a	-0.8	0.5
Downward shortwave (SW)	M	0.3 [-]	24 h	0.5°	-0.8	n/a	-0.5
Downward longwave (LW)	A	20 W/m^2	24 h	0.5°	0.5	-0.5	n/a
Catchment deficit (catdef)	A	0.24 kg/m^2/h	3 h	0.5°			
Surface excess (srfexc)	A	0.16 kg/m^2/h	3 h	0.5°			

TABLE 3. Overview of experiments.

Experiment	No CPCU precipitation corrections	With CPCU precipitation corrections
Ensemble open loop (model-only)	OL	OL_P
Assimilation	DA	DA_P

TABLE 4. Skill metrics at SCAN, USCRN, SMOSMANIA and Oznet sites, with indication of the half value of the 95% confidence intervals (CI) averaged over all experiments (for asymmetric confidence intervals, the mean half value is shown). The metrics are calculated using all 3-hourly time steps during the warm season across 4 years. N is the number of sites for each network, with the number of clusters between parentheses. Clusters are not used for the Oznet and SMOSMANIA sites, where the CIs are conservatively estimated as the simple average of the CIs over individual sites. Temperature validation is not performed at the Oznet sites for lack of sufficient data pairs at each individual 3-hour time interval.

ubRMSD	N	OL	OL_P	DA	DA_P	CI
Surface soil moisture [m^3/m^3]						
SCAN	62(20)	0.062	0.057	0.056	0.052	±0.003
USCRN	41(21)	0.061	0.056	0.055	0.052	±0.004
SMOSMANIA	9	0.053	0.050	0.049	0.046	±0.014
Oznet	38	0.058	0.055	0.053	0.052	±0.038
Root-zone soil moisture [m^3/m^3]						
SCAN	62(20)	0.043	0.038	0.039	0.037	±0.009
USCRN	41(21)	0.045	0.040	0.041	0.039	±0.005
SMOSMANIA	9	0.044	0.041	0.041	0.040	±0.032
Oznet	17	0.053	0.049	0.048	0.045	±0.053
Surface soil temperature [K]						
SCAN	82(21)	2.02	1.98	2.02	2.00	±0.03
USCRN	48(23)	1.85	1.80	1.87	1.83	±0.03
SMOSMANIA	9	1.83	1.81	1.83	1.81	±0.13
anomR	N	OL	OL_P	DA	DA_P	CI
Surface soil moisture [-]						
SCAN	59(18)	0.48	0.58	0.65	0.69	±0.04
USCRN	41(21)	0.54	0.63	0.67	0.72	±0.03
SMOSMANIA	9	0.56	0.59	0.63	0.67	±0.15
Oznet	29	0.57	0.62	0.67	0.69	±0.27
Root-zone soil moisture [-]						
SCAN	57(18)	0.57	0.66	0.69	0.70	±0.07
USCRN	41(21)	0.62	0.69	0.68	0.70	±0.07
SMOSMANIA	9	0.60	0.63	0.62	0.63	±0.35
Oznet	7	0.62	0.61	0.71	0.77	±0.45
Surface soil temperature [-]						
SCAN	79(20)	0.78	0.79	0.79	0.79	±0.005
USCRN	48(23)	0.78	0.80	0.80	0.80	±0.004
SMOSMANIA	9	0.86	0.86	0.87	0.87	±0.016

TABLE 5. Spatial average of the long-term and seasonal statistics for experiment OL_P and differences with experiment DA_P for various land surface variables. OL_P statistics include the temporal mean (mean OL_P) and standard deviation (stdv OL_P). Difference statistics include the difference in the mean, mean of the absolute differences, difference in the standard deviation (MDiff, MabsDiff, SDiff; DA_P minus OL_P) and the ubRMSD between DA_P and OL_P results. Statistics are calculated as the spatial average of the metrics for North America (20-50^o N) during the period 1 July 2010 - 1 July 2014.

Metric	All months	DJF	MAM	JJA	SON
	surface soil moisture [m^3/m^3]				
mean OL_P	0.21	0.22	0.22	0.20	0.19
stdv OL_P	0.042	0.027	0.027	0.040	0.033
MDiff	8.24E-04	4.84E-04	2.15E-04	1.06E-03	1.94E-03
MabsDiff	2.55E-03	3.66E-03	2.71E-03	3.40E-03	3.66E-03
SDiff	4.20E-03	2.55E-03	5.01E-03	4.99E-03	6.78E-03
ubRMSD	0.019	0.015	0.016	0.020	0.022
	root-zone soil moisture [m^3/m^3]				
mean OL_P	0.21	0.21	0.22	0.20	0.19
stdv OL_P	0.033	0.023	0.022	0.029	0.022
MDiff	9.78E-04	1.82E-04	3.70E-04	1.70E-03	2.18E-03
MabsDiff	2.65E-03	3.40E-03	2.70E-03	3.72E-03	3.73E-03
SDiff	3.69E-03	1.40E-03	4.28E-03	4.94E-03	5.54E-03
ubRMSD	0.016	0.013	0.014	0.017	0.017
	surface soil temperature [K]				
mean OL_P	286.05	275.97	285.04	296.23	289.88
stdv OL_P	9.05	3.84	6.14	4.02	5.30
MDiff	-1.33E-02	-4.10E-03	-3.69E-03	-2.13E-02	-3.34E-02
MabsDiff	3.14E-02	3.12E-02	3.42E-02	6.17E-02	5.45E-02
SDiff	2.27E-03	3.37E-03	1.34E-02	3.53E-02	2.22E-02
ubRMSD	0.33	0.22	0.27	0.40	0.35
	sensible heat flux [W/m^2]				
mean OL_P	47.63	16.96	62.22	72.67	46.57
stdv OL_P	99.44	58.39	112.72	113.09	93.28
MDiff	-0.37	-0.08	-0.07	-0.65	-0.96
MabsDiff	0.83	0.43	0.92	1.94	1.43
SDiff	0.14	-0.02	0.42	0.37	-0.33
ubRMSD	11.35	3.83	9.94	15.81	11.71
	latent heat flux [W/m^2]				
mean OL_P	46.72	18.57	53.67	76.48	44.90
stdv OL_P	67.05	27.59	65.88	85.06	56.17
MDiff	0.49	0.11	0.11	0.85	1.23
MabsDiff	1.01	0.53	1.11	2.34	1.77
SDiff	2.17	0.42	1.52	3.19	3.54
ubRMSD	13.18	4.12	10.98	18.42	13.88
	total runoff [mm/d]				
mean OL_P	0.35	0.32	0.51	0.32	0.28
stdv OL_P	1.48	1.22	1.72	1.15	1.10
MDiff	-4.32E-04	-9.42E-03	-1.10E-03	6.48E-03	4.53E-03
MabsDiff	1.37E-02	1.52E-02	1.88E-02	1.92E-02	1.77E-02
SDiff	2.23E-02	-2.56E-02	2.25E-02	3.46E-02	3.89E-02
ubRMSD	0.27	0.18	0.29	0.24	0.25

List of Figures

4 Change in anomaly time series correlation (ΔanomR) due to data assimilation (DA, without CPCU correction) at (circles) SCAN and (triangles) USCRN sites for (a) surface and (b) root-zone soil moisture. Statistically significant changes are marked by larger symbols. Metrics are calculated across 3-hourly time steps during the period 1 July 2010 - 1 July 2014. N is the number of sites and the titles indicate the spatial mean ΔanomR across all sites with clustering. All sites are located in areas with limited vegetation and topographic complexity based on model parameters, indicated by the green background shading. 56

5 Same as Figure 4b but for ΔubRMSD in root-zone soil moisture only and including sites outside the areas with favorable conditions indicated by the green shading. Metrics are for (a) SCAN (circles) and USCRN (triangles), (b) Oznet, and (c) SMOSMANIA sites. 57

6 Performance of various open loop (OL, OL_P) and data assimilation (DA, DA_P) experiments in terms of anomR at individual reference grid cells (Table 1) for (a) surface soil moisture, (b) root-zone soil moisture and (c) soil temperature. The metrics are calculated across all analysis and forecast time steps during the period 1 July 2010 - 1 July 2014, except for temperature where the anomR is calculated for warm season across four years. 'All' refers to the average skill across all available reference grid cells. 58

7 Performance of various open loop (OL, OL_P) and data assimilation (DA, DA_P) experiments in terms of (a,d) ubRMSD (b,e) bias and (c,f) R for (top) surface and (bottom) root-zone soil moisture at individual reference grid cells (Table 1). The metrics are calculated across all analysis and forecast time steps during the period 1 July 2010 - 1 July 2014. 'All' refers to the average skill across all available reference grid cells. 59

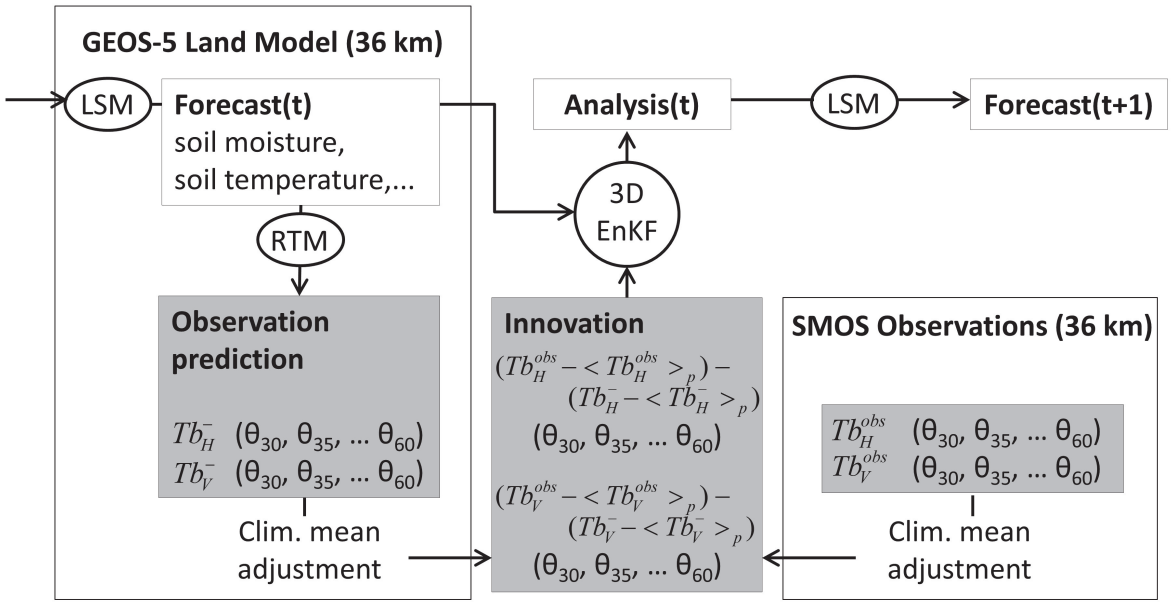

FIG. 1. Schematic of the multi-angular $(\theta_{30}, \theta_{35}, \ldots, \theta_{60})$ and multi-polarization (H, V) SMOS Tb assimilation scheme, using a 3D-EnKF and mean climatological rescaling to address biases. LSM is the land surface model, RTM is the L-band radiative transfer model.

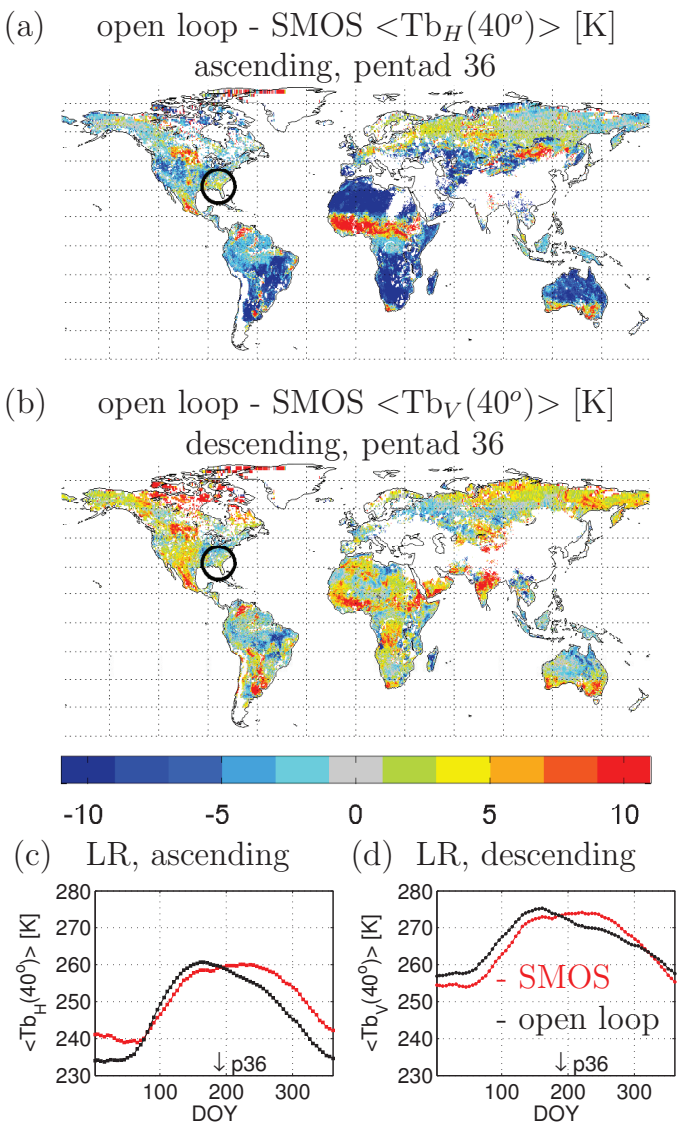

FIG. 2. (a) Difference between climatological mean Tb(40°) forecasts (ensemble mean open loop without CPCU precipitation corrections) and SMOS observations for pentad 36, H-polarization and ascending orbits. (b) Same as (a), but for V-polarization and descending orbits. The center of the circle marks the location of the Little River watershed. (c) Time series of (black dots) climatological mean Tb(40°) forecasts (ensemble mean open loop without CPCU precipitation corrections) and (red dots) SMOS observations at Little River (LR), Georgia, for H-polarization and ascending orbits. The arrow indicates pentad 36, corresponding to the pentad for subplots a and b. (d) Same as (c) but for V-polarization and descending orbits.

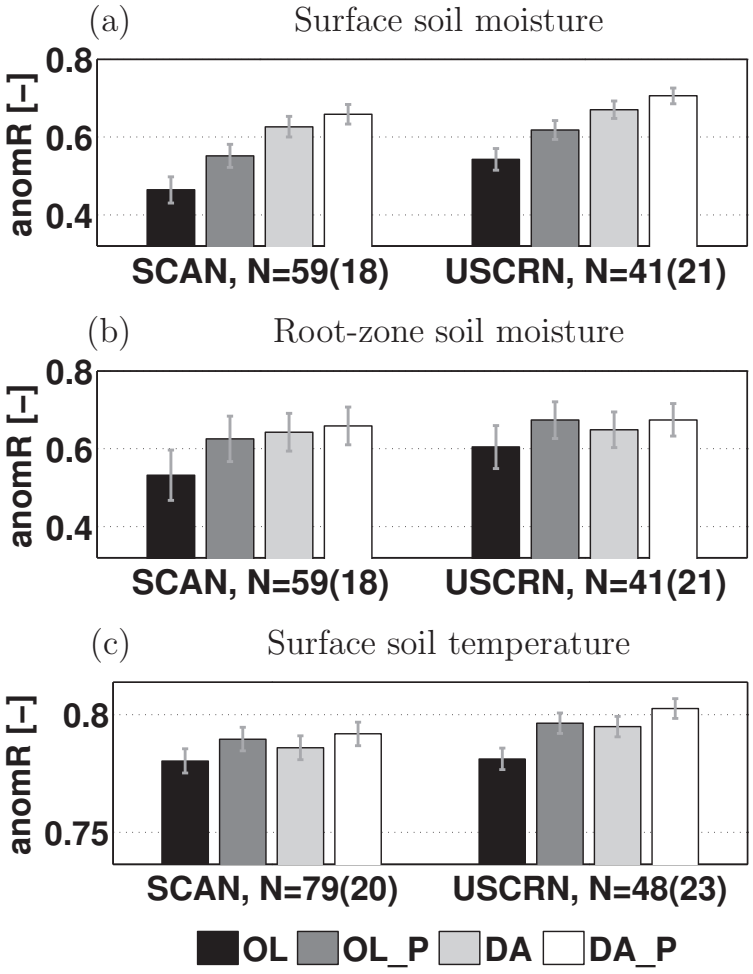

FIG. 3. Anomaly time series correlation coefficient (anomR) for two open loop experiments (OL, OL_P) and two data assimilation experiments (DA, DA_P), averaged across individual SCAN and USCRN sites for (a) surface soil moisture, (b) root-zone soil moisture and (c) surface soil temperature. The metric is calculated for the period 1 July 2010 - 1 July 2014, except for temperature where the anomR is calculated for the warm season only (details in section 4b). N is the number of sites for each network, with the number of clusters in parentheses.

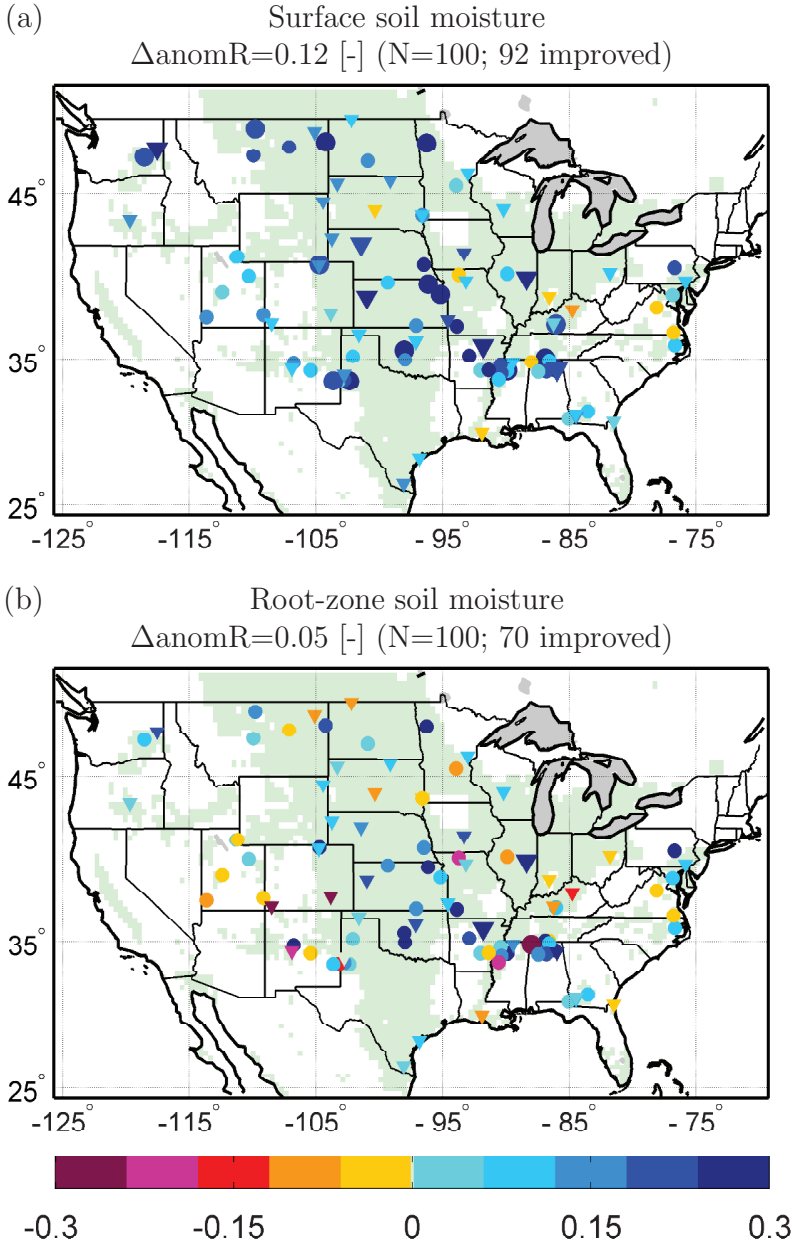

FIG. 4. Change in anomaly time series correlation (ΔanomR) due to data assimilation (DA, without CPCU correction) at (circles) SCAN and (triangles) USCRN sites for (a) surface and (b) root-zone soil moisture. Statistically significant changes are marked by larger symbols. Metrics are calculated across 3-hourly time steps during the period 1 July 2010 - 1 July 2014. N is the number of sites and the titles indicate the spatial mean ΔanomR across all sites with clustering. All sites are located in areas with limited vegetation and topographic complexity based on model parameters, indicated by the green background shading.

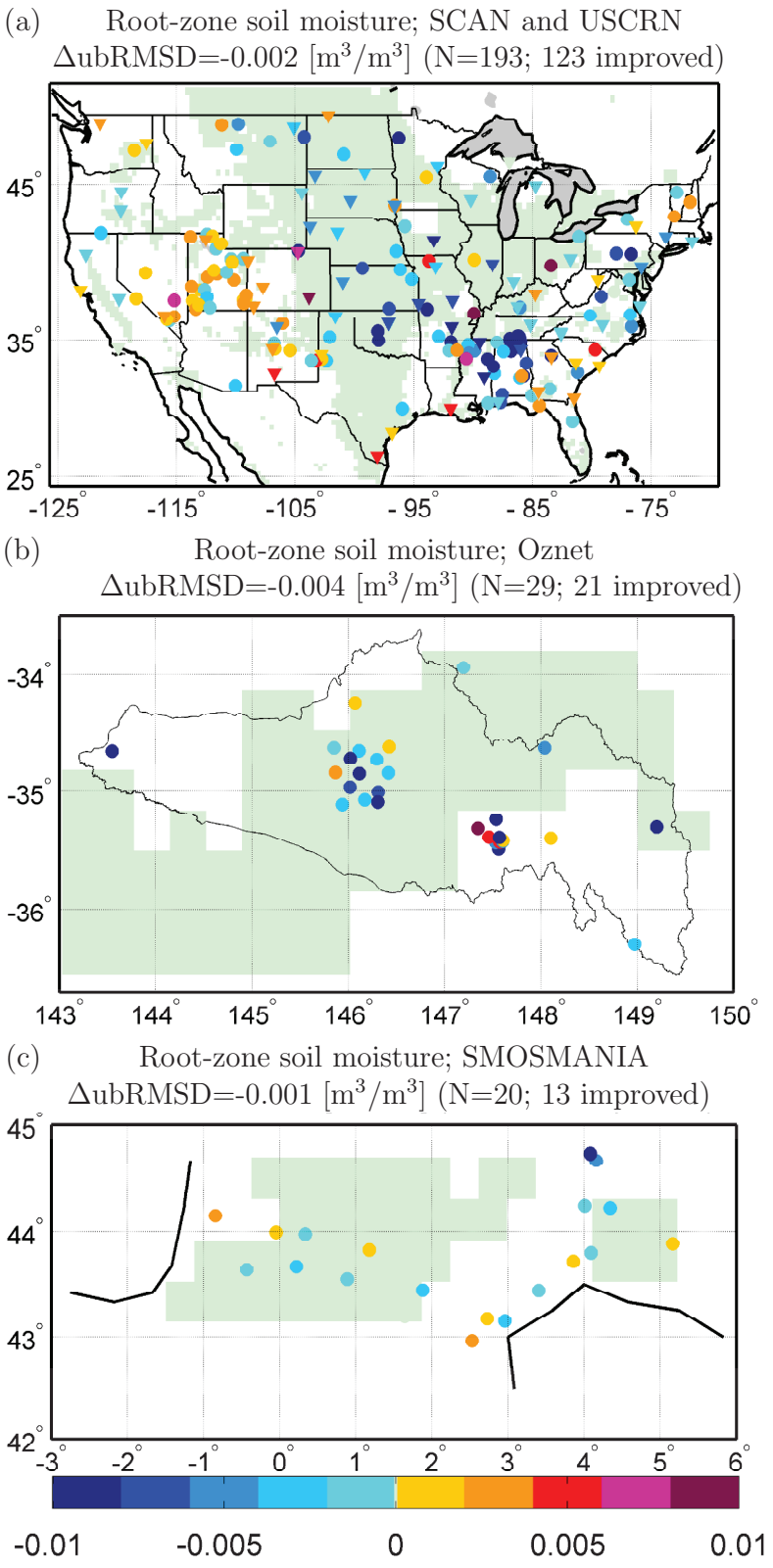

FIG. 5. Same as Figure 4b but for ΔubRMSD in root-zone soil moisture only and including sites outside the areas with favorable conditions indicated by the green shading. Metrics are for (a) SCAN (circles) and USCRN (triangles), (b) Oznet, and (c) SMOSMANIA sites.

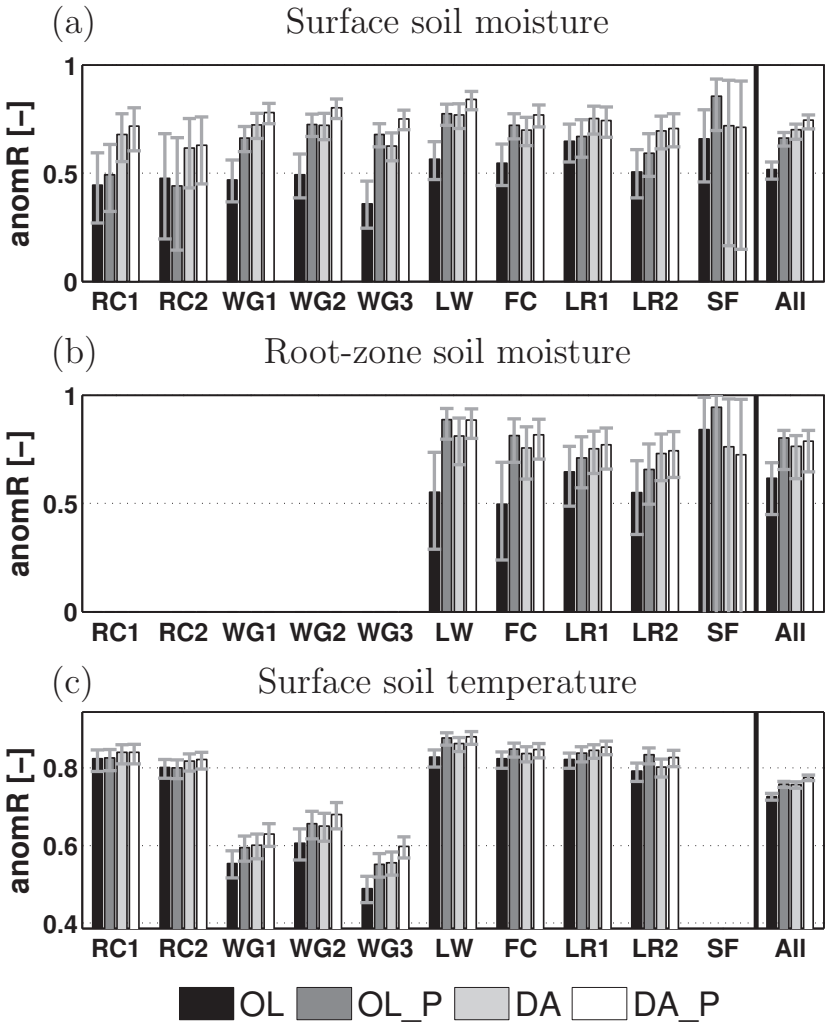

FIG. 6. Performance of various open loop (OL, OL_P) and data assimilation (DA, DA_P) experiments in terms of anomR at individual reference grid cells (Table 1) for (a) surface soil moisture, (b) root-zone soil moisture and (c) soil temperature. The metrics are calculated across all analysis and forecast time steps during the period 1 July 2010 - 1 July 2014, except for temperature where the anomR is calculated for warm season across four years. 'All' refers to the average skill across all available reference grid cells.

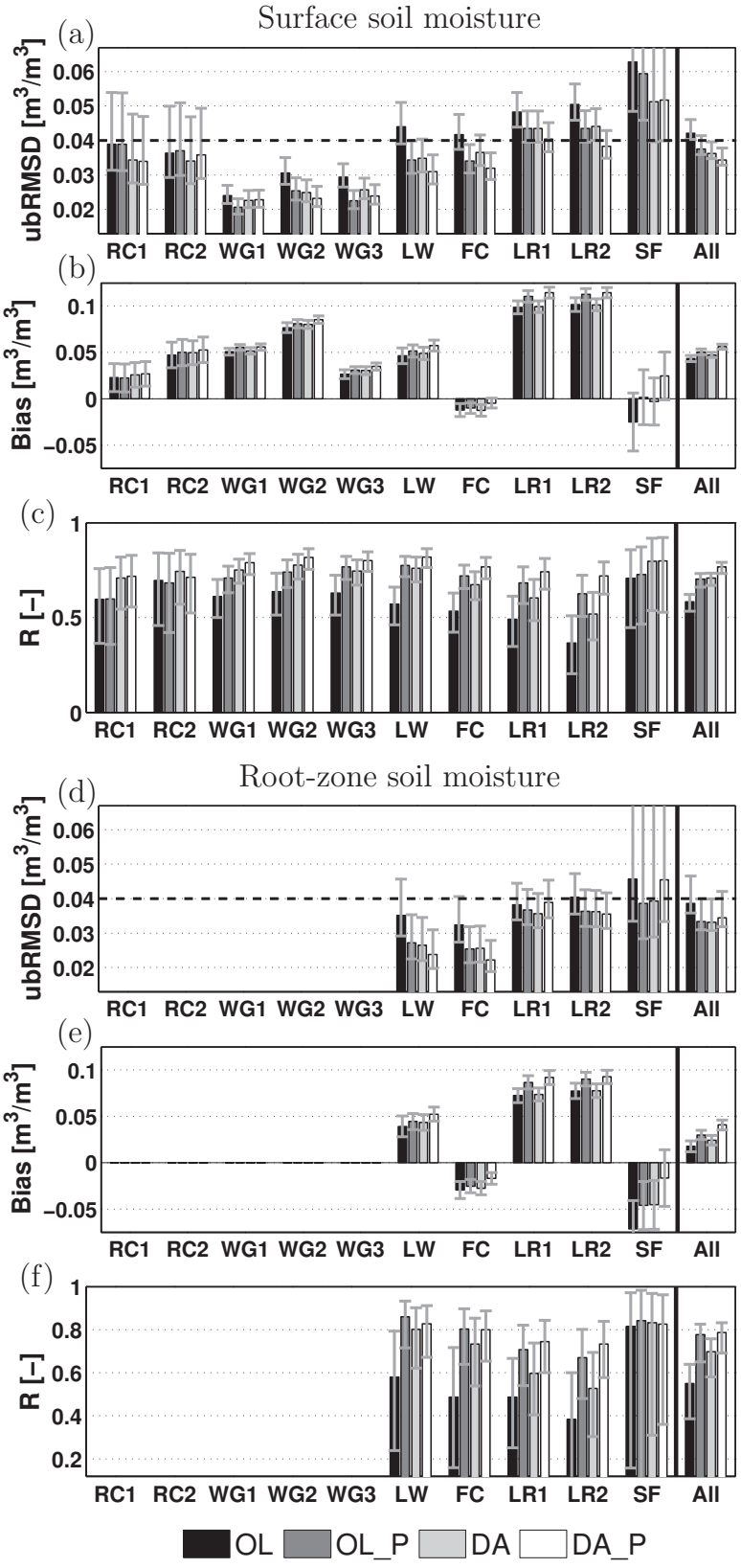

FIG. 7. Performance of various open loop (OL, OL_P) and data assimilation (DA, DA_P) experiments in terms of (a,d) ubRMSD (b,e) bias and (c,f) R for (top) surface and (bottom) root-zone soil moisture at individual reference grid cells (Table 1). The metrics are calculated across all analysis and forecast time steps during the period 1 July 2010 - 1 July 2014. 'All' refers to the average skill across all available reference grid cells.

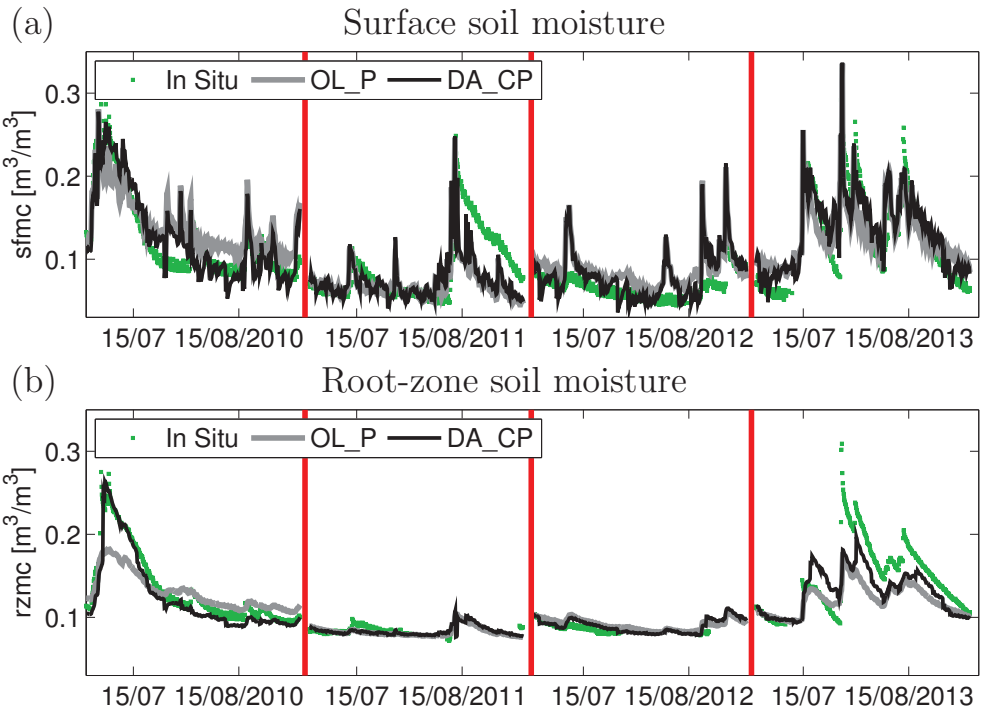

FIG. 8. Soil moisture time series at the Fort Cobb (Oklahoma, 35.34 ^{o}N, 98.40 ^{o}W) reference grid cell during each July and August in the 4-year experiment period. Time series of (a) in situ observed and modeled surface soil moisture (sfmc) for the open loop (OL_P) and assimilation integration (DA_P); (b) same as (a) but for root-zone soil moisture (rzmc).

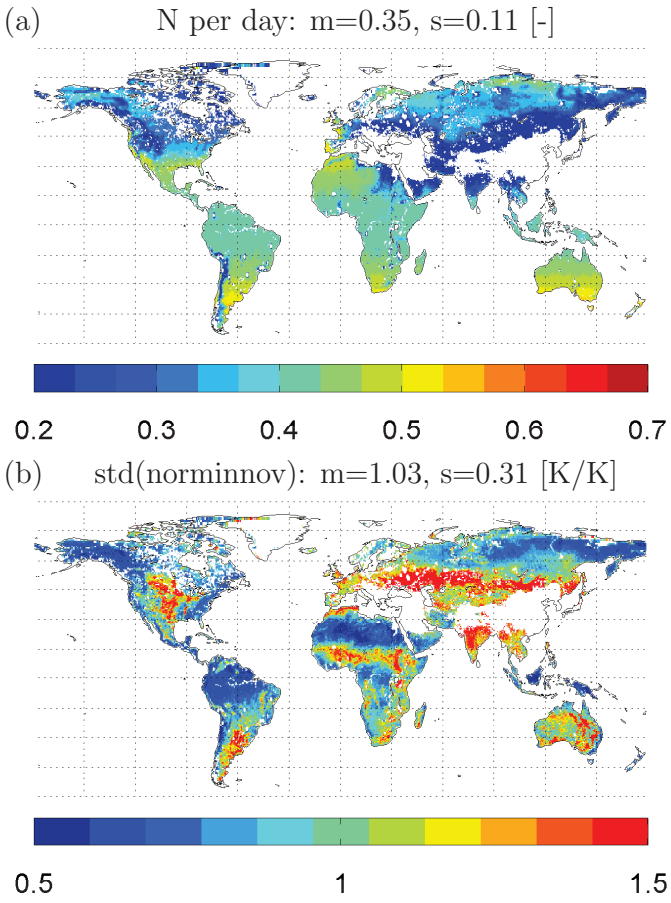

FIG. 9. Statistics of the observation-minus-forecast residuals (innovations) for experiment DA_P, calculated for 1 July 2010 - 1 July 2014. (a) Average number (N) of assimilated Tb sets per day; (b) standard deviation of normalized Tb innovations (norminnov), averaged across both polarizations and all available incidence angles. The titles indicate the spatial mean (m) and standard deviation (s) across each map.

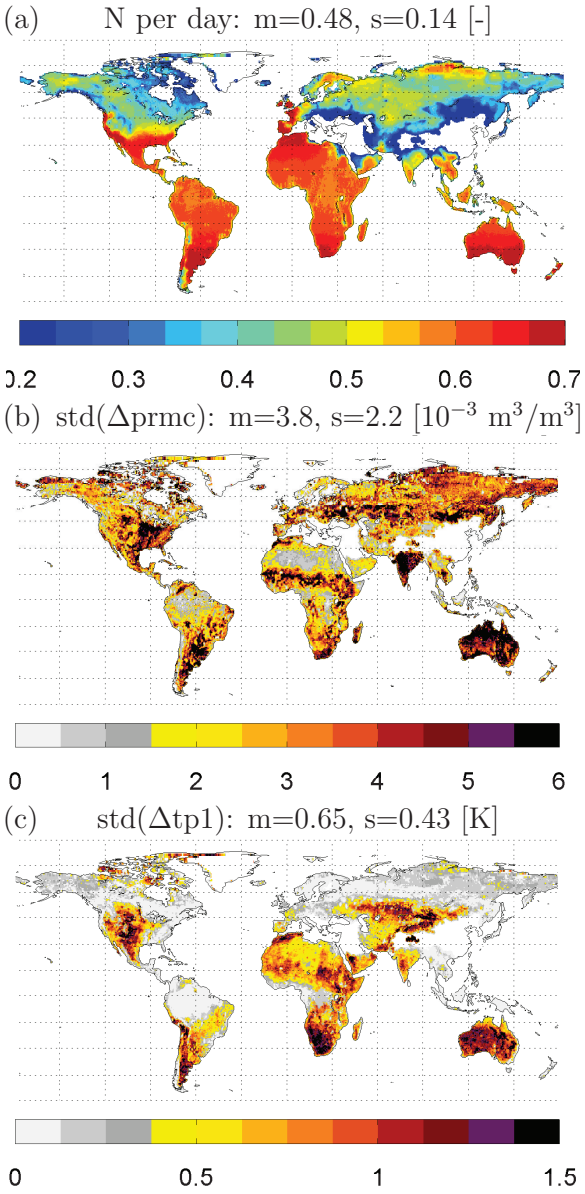

FIG. 10. Statistics of the analysis increments for experiment DA_P, calculated for 1 July 2010 - 1 July 2014. (a) Average number (N) of increments per day; (b) standard deviation in soil column water increments (Δprmc); (c) standard deviation in soil temperature increments (Δtp1). The titles indicate the spatial mean (m) and standard deviation (s) across each map.

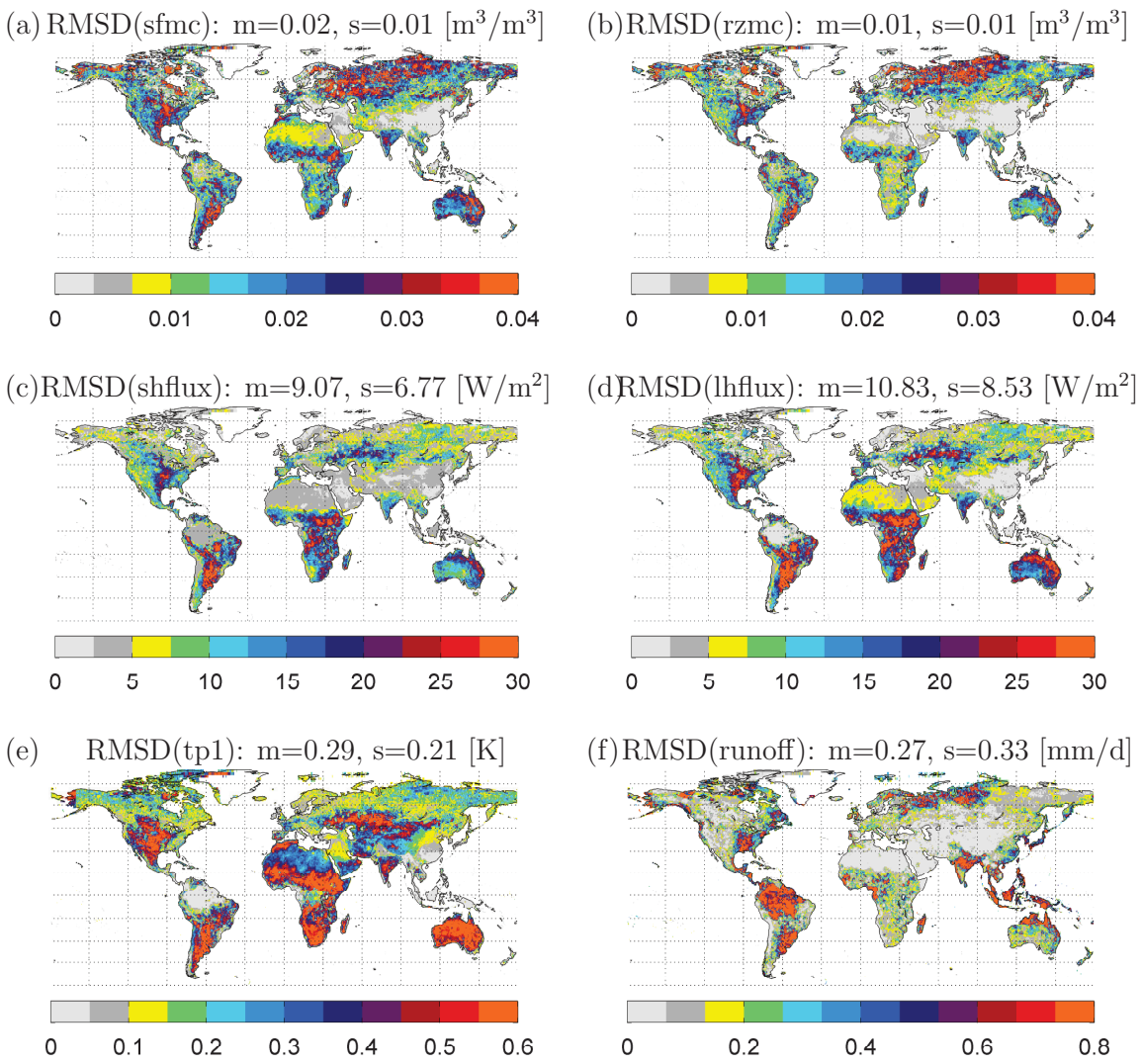

FIG. 11. Temporal RMSD between the open loop simulation (OL_P) and the assimilation (DA_P) estimates for various land surface variables: (a) surface soil moisture (sfmc), (b) root-zone soil moisture (rzmc), (c) sensible heat flux (shflux), (d) latent heat flux (lhflux), (e) soil temperature (tp1), and (f) total runoff. The RMSD is calculated using 3-hourly model output during 1 July 2010 - 1 July 2014. The titles indicate the spatial mean (m) and standard deviation (s) across each map.

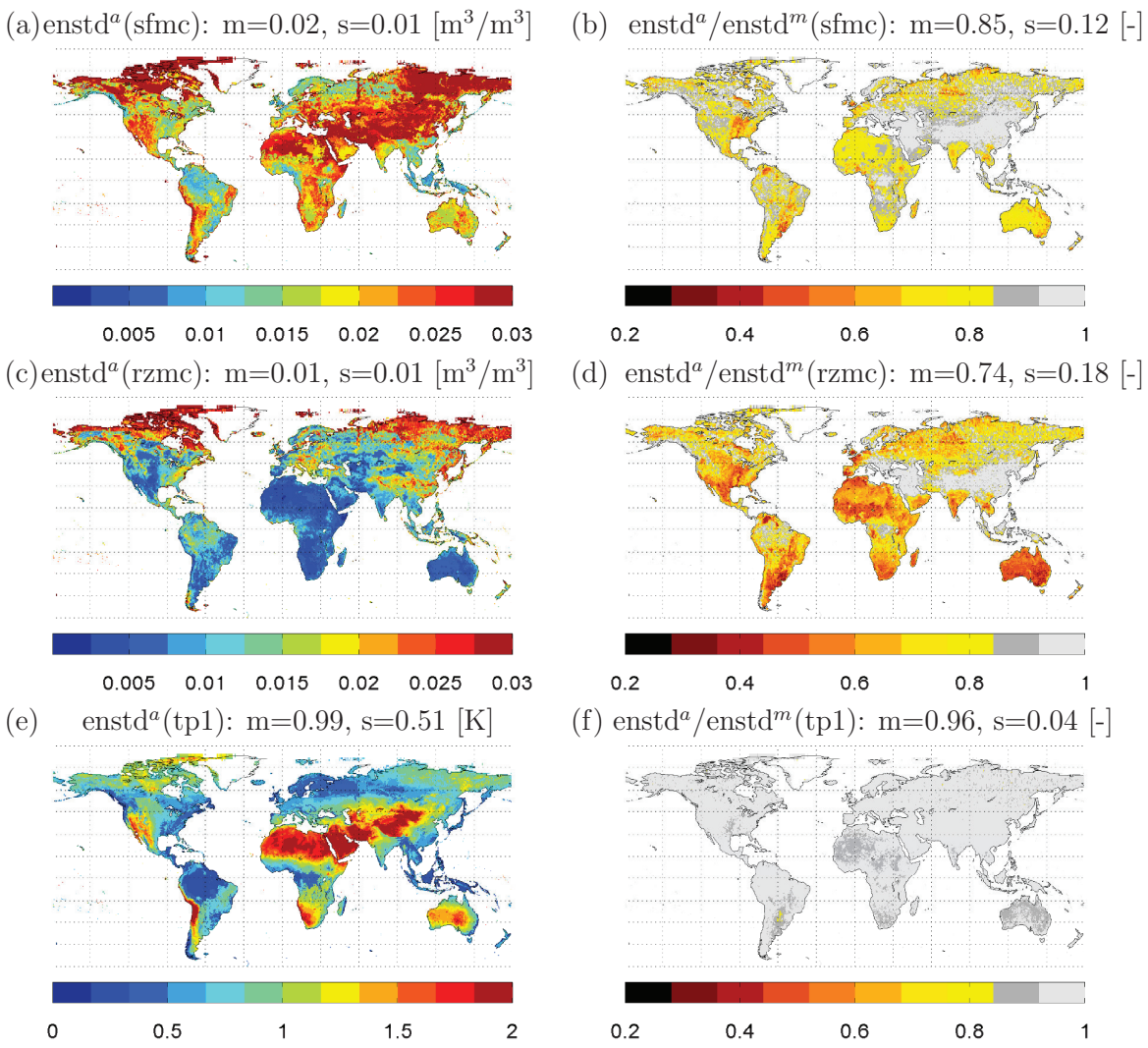

FIG. 12. (Left) temporal mean ensemble error standard deviation (enstd^a) for assimilation case DA_P and (right) ratio of ensemble error standard deviations (enstd^a/enstd^m) in the assimilation experiment DA_P versus the model-only experiment OL_P, for (a-b) surface soil moisture (sfmc), (c-d) root-zone soil moisture (rzmc), and (e-f) soil temperature (tp1). The values are based on 3-hourly estimates and averaged over the period 1 July 2010 - 1 July 2014. The titles indicate the spatial mean (m) and standard deviation (s) across each map.